

Vincenzo Lombardo
Alberto Zazzaro
(Eds)

Annals of CRISEI 2025

**Inequality, Access and Mobility: Evidence from
Education, Credit and Migration**

CRISEI

ISBN: 978-88-98279-16-6

Contents

Introduction	IV
Vincenzo Lombardo and Alberto Zazzaro	
 Does Your Person Hold a Passport? The Identikit of Italian Emigrants in the Age of Mass Migration	1
Maria Rosaria Carillo and Salvatore Carrozzo	
 Logged Out: How COVID-19 School Closures Deepened Educational Poverty in Italy	24
Emma Sarno	
 Discrimination of Immigrants in Mortgage Pricing and Approval: Evidence from Italy	48
Paolo Emilio Mistrulli, Md Taslim Uddin and Alberto Zazzaro	
 The Geopolitics of Aid: U.S. and China's Strategic Financial Flows	111
Emmanuel Caiazzo, Pietro Panizza and Alberto Zazzaro	
 Authority, Information, and Credit Terms: Evidence from Small Business Lending	140
Andrea Bellucci, Alexander Borisov and Alberto Zazzaro	
 Status Seeking and Positional Consumption Under Mobility Restrictions During COVID-19	173
Vincenzo Lombardo and Angela Pepe	

Contributors

Andrea Bellucci, University of Insubria

Alexander Borisov, University of Cincinnati

Maria Rosaria Carillo, University of Naples Parthenope

Emmanuel Caiazzo, University of Naples Parthenope

Salvatore Carrozzo, University of Naples Parthenope

Paolo Emilio Mistrulli, Bank of Italy

Vincenzo Lombardo, University of Naples Parthenope

Pietro Panizza, University of Calabria

Angela Pepe, ISTAT

Emma Sarno, University of Naples L'Orientale

Md Taslim Uddin, Asian University for Women

Alberto Zazzaro, University of Naples Federico II

Introduction

Vincenzo Lombardo

University of Naples Parthenopoe

`vincenzo.lombardo@uniparthenope.it`

Alberto Zazzaro

University of Naples Federico II, CSEF, and MoFiR

`alberto.zazzaro@unina.it`

The present volume, *Inequality, Access and Mobility: Evidence from Education, Credit and Migration*, brings together six contributions that share a common concern with the mechanisms through which economic, social, and territorial inequalities translate into unequal access to resources and heterogeneous patterns of mobility. Although the chapters span different historical periods, levels of analysis, and areas of inquiry, they collectively show that opportunities depend not only on the formal availability of markets, services, or institutions, but also on access costs, the distribution of information, social networks, and individuals' positions within broader structures of power. Migration, learning, access to credit, the receipt of international aid, and the display of social status through consumption are all shaped by constraints whose effects are unevenly distributed across social groups.

The common thread running through the volume is therefore the relationship between mobility and inequality. Mobility is understood in a broad sense: geographical, as in the movements of Italian emigrants; educational and social, as reflected in learning opportunities during the pandemic; economic, as in access to mortgages and business finance; geopolitical, as in competition for influence through international aid; and relational and symbolic, as in the pursuit of status through consumption. Across all the contributions, access emerges as a selective process mediated by public policies and social restrictions, technologies, organizational structures, forms of knowledge, and external shocks. Institutions can reduce costs and broaden opportunities, but they can also reproduce existing disparities or generate new ones. The volume therefore invites reflection on the complex processes that shape access to opportunities and resources, as well as the choices made by individuals and institutions. These processes determine which individuals and social groups are able to participate in different forms of mobility and which instead risk being excluded, unable to overcome the socioeconomic barriers that hinder equal access to opportunities for advancement and to the resources required to pursue them.

The first chapter, *Does Your Person Hold a Passport? The Identikit of Italian Emigrants in the Age of Mass Migration*, by Maria Rosaria Carillo and Salvatore Carrozzo, addresses two classic questions in the economics of migration: the selection of those who migrate and the

sorting of migrants across destinations. The study draws on an original administrative source consisting of passport applications preserved in the State Archive of Naples to reconstruct the profile of Italian emigrants between the late nineteenth and early twentieth centuries. The distinction between a more restrictive regulatory phase, from 1888 to 1900, and the subsequent liberal period makes it possible to assess how declining migration costs and changes in the regulatory framework affected emigration flows. The findings indicate that the easing of restrictions led to higher emigration from poorer neighborhoods, while leaving the occupational composition of migrants largely unchanged. The chapter thus provides new micro-level evidence on emigration from Southern Europe and on the processes of migrant selection and sorting across destinations.

In the second chapter, *Logged Out: How COVID-19 School Closures Deepened Educational Poverty in Italy*, Emma Sarno shifts the focus to educational mobility and the distributional consequences of the sudden transition to remote learning. The contribution distinguishes among the pre-pandemic period, the 2020–2021 pandemic shock, and the outcomes observed in the post-pandemic period, combining micro- and macro-level evidence on socioeconomic conditions, territorial disparities, digital access, and educational achievement. The analysis shows that remote learning was far from a neutral emergency response. Unequal access to devices, reliable internet connections, and adequate home learning environments turned the digital divide into a mechanism that amplified pre-existing educational inequalities. The effects were particularly severe for students from disadvantaged households and from the southern regions of Italy. The pandemic thus brought into sharper relief a structural mechanism through which technological inequality translates into more limited learning opportunities and deeper forms of educational poverty.

The third chapter, *Discrimination of Immigrants in Mortgage Pricing and Approval: Evidence from Italy*, by Paolo Emilio Mistrulli, Md Taslim Uddin, and Alberto Zazzaro, examines disparities between immigrants and natives in the Italian mortgage market. Combining two datasets from the Italian Credit Register—covering mortgage contracts and banks’ preliminary information requests on prospective borrowers—with data from the Bank of Italy’s Survey on Household Income and Wealth, the authors investigate both the pricing of credit and the probability of mortgage approval. Immigrants are found to pay higher interest rates: the differential ranges from 20 to 24 basis points for single-name mortgages and from 28 to 40 basis points for jointly held mortgages. In addition, immigrants face a 2.7 percentage-point lower probability of obtaining mortgage approval. The gap narrows, although it does not disappear, when borrowers have longer credit histories or when they deal with cooperative banks. The evidence suggests that a substantial part of the disparity reflects the informational difficulties faced by financial intermediaries when assessing culturally distant

borrowers, while also pointing to the risk of persistent differences in treatment.

The fourth chapter extends the question of access to resources to the sphere of international relations. In *The Geopolitics of Aid: U.S. and China’s Strategic Financial Flows*, Emmanuel Caiazzo, Pietro Panizza, and Alberto Zazzaro examine whether development aid has become an instrument of strategic competition between the United States and China. Using data from the OECD Creditor Reporting System and the Global Chinese Development Finance Dataset for the period 2000–2019, the study employs an instrumental-variable approach to investigate the allocation of financial commitments across recipient countries. The results show that the United States increases its aid commitments in countries where Chinese financing expands. The U.S. response is stronger when political or commercial ties with recipient countries are more substantial and becomes more pronounced following Xi Jinping’s rise to power. International aid thus emerges not only as a development instrument but also as a strategic lever through which major powers seek to counter competitors’ influence and consolidate political and economic relationships.

The fifth contribution, *Authority, Information, and Credit Terms: Evidence from Small Business Lending*, by Andrea Bellucci, Alexander Borisov, and Alberto Zazzaro, returns to the credit market to investigate how banks’ organizational architecture shapes lending conditions for small businesses. The study exploits variation generated by the organizational restructuring of an Italian bank and examines the interaction between decision-making authority and the production of information about borrowers. When the overlap between those who gather information and those who make lending decisions increases, the use of collateral declines, while interest rates remain unchanged. The effect is stronger when loan officers have access to qualitative information and tacit local knowledge about borrowers, or when their effective decision-making authority is limited. The chapter therefore shows that access to credit depends not only on firms’ characteristics but also on the way information and decision-making power are allocated within financial intermediaries.

The sixth chapter, *Status Seeking and Positional Consumption under Mobility Restrictions during COVID-19*, by Vincenzo Lombardo and Angela Pepe, examines the social visibility of consumption and its role as a signal of relative status. Combining microdata from the Italian Household Budget Survey with municipality-level information on excess mortality, the authors study the effect of mobility restrictions on expenditures for visible and positional goods. The decline in such consumption during periods of reduced social interaction confirms that its value depends in part on the presence of an audience and on opportunities for social observation. The contribution provides new evidence on the behavioral nature of conspicuous consumption and on the relationship between visibility, status, and social networks, showing how restrictions on physical mobility can alter the channels through which

individuals communicate their relative social position.

Taken together, the six chapters provide a multifaceted account of the processes linking inequality, access, and mobility. Barriers may take the form of administrative costs, technological shortcomings, information asymmetries, cultural distance, organizational arrangements, or restrictions on movement; yet their effects ultimately converge in determining who is able to translate resources and aspirations into effective opportunities. The contributions also show that these mechanisms are not immutable. Migration policies, investments in digital infrastructure, banking models, the allocation of decision-making authority, and international strategies can all alter both their intensity and their direction. By combining historical and contemporary perspectives and moving across micro- and macro-levels of analysis, the volume highlights how mobility can be both an outcome of inequality and a channel through which inequality is reproduced or reduced. It is precisely this tension that gives coherence to the collection and makes it relevant to understanding the multiple forms that exclusion and opportunity assume in modern economies.

Does Your Person Hold a Passport? The Identikit of Italian Emigrants in the Age of Mass Migration*

Maria Rosaria Carillo[†]

Salvatore Carrozzo[‡]

Abstract

This paper examines the selection and sorting of Italian emigrants during the Age of Mass Migration (late 19th to early 20th centuries) using a novel dataset of passport applications from the State Archive of Naples. It addresses key questions in migration economics: who migrates (selection) and where they go (sorting). The analysis covers two policy periods stringent (1888–1900) and liberal (1901 onward) allowing us to test how migration costs and policy shifts influenced emigrant profiles. Our findings reveal that reduced migration costs increased emigration from poorer neighborhoods, but occupational composition remained stable. The paper contributes to the literature by introducing a rich, previously untapped administrative source, offering micro-level insights into a Southern European emigration context, and comparing historical sorting patterns across destinations.

JEL Classification codes: F22, O15, J61

Keywords: Immigration, Selection, Sorting

*Maria Rosaria Carillo and Salvatore Carrozzo greatly acknowledge research funds received by the European Union – Next Generation EU, under the GRINS project – Growing Resilient, Inclusive and Sustainable (GRINS PE00000018 – CUP E63C22002140007) “The Persistence effects of international Migration through the channels of human capital and non-cognitive skills”. Authors also thanks the National State Archive of Naples for their assistance in this research.

[†]Department of Business and Economics (DISAE), University of Naples "Parthenope". Email: mr.carillo@uniparthenope.it.

[‡]Department of Business and Economics (DISAE), University of Naples "Parthenope". Email: salvatore.carrozzo@collaboratore.uniparthenope.it.

1 Introduction

Today, migrants constitute approximately 3.7% of the global population—over 280 million individuals—a figure that continues to grow steadily, driven by disparities in economic opportunity, political instability, and the shrinking costs of mobility (UN DESA, 2025). The movement of people across borders generates profound transformations, reshaping labor markets, altering demographic structures, transmitting cultural practices, and challenging political institutions in both sending and receiving societies. Understanding the dynamics of who migrates (selection) and where they go (sorting) is a pressing concern for policymakers tasked with managing integration, economic planning, and social cohesion. However, the lack of information on the characteristics of emigrants recorded in the country of origin does not allow to characterize the selection of migrants with respect to the country of origin and the sorting of emigrants across destinations.

This study addresses the source-country data gap by introducing and analyzing a novel, rich, and previously untapped source: administrative records of passport applications from late 19th and early 20th century Italy. We have collected individual characteristics of passport applicants over 40,000 individual application forms from the State Archive of Naples, spanning the pivotal years 1888 to 1909. This period, known as the Age of Mass Migration, witnessed the exodus of over 14 million Italians, an important migration event for Italy and for the demographics of the destination countries. These application forms, filled out by police officials, constitute a unique origin-side, pre-migration dataset. They record individual characteristics (age, civic status, gender, occupation, literacy, place of birth), wealth characteristics (criminal records, physical characteristics) and migration characteristics (intended destination).

This unique data allows us to directly observe the pool of potential emigrants at the moment they declared their intent to leave. We can therefore measure selection by comparing applicants

to the average population characteristics, and analyze sorting by linking individual profiles to declared destinations. Furthermore, our time span captures two different migration policies: between 1888-1901 the Italian government has hindered the emigration of male Italians abroad by excluding those eligible for the military service, while the 1901 Italian Emigration Law shifted dramatically to actively encourage and facilitate emigration, as remittances became a vital source of foreign capital and a social safety valve. This policy change allows us to test how a significant reduction in institutional barriers and migration costs affected both the type of person seeking to emigrate (selection) and their choice of destination (sorting).

We first provide descriptive evidences on the characteristics of passport applicants to depict who was used to travel temporarily or permanently. Our dataset includes individual characteristics of more than 40,000 individuals who applied for a passport at Naples police headquarters between 1888 and 1909. Since the passport lasted only few years, we observe the renewal of the passport for those who stayed permanently abroad or for those traveling several times around the world. Although we cannot distinguish between international merchants and proper emigrants, we use several individual dimensions to understand whether people were emigrating to spend some years abroad. Once depicted the identikit of the potential leavers, we plot the average characteristics of emigrants across several destinations.

Our findings reveal a consistent pattern of positive selection. Passport applicants were significantly more literate and held higher-status occupations than the average Southern Italian. This positive selection persisted even after the 1901 policy liberalization, which spurred a dramatic increase in applications from poorer urban neighborhoods in Naples but did not change the positive selection of passport applicants measured as occupational status. Regarding sorting, we find strong evidence of skill-based destination choice. Literate individuals and skilled artisans were significantly more likely to target the industrializing United States, while

agricultural laborers and those following family networks were more likely to choose Argentina. This indicates that economic rationality and social networks operated simultaneously to direct the Italian diaspora along predictable channels.

This paper makes three primary contributions to the literature. First, it introduces and validates a major new historical microdataset, demonstrating the unparalleled value of passport applications for migration research. This source bridges the gap between origin and destination records, providing the detailed, individual-level information needed to resolve long-standing debates on selection and sorting. The economics of migration has long grappled with the fundamental puzzle of migrant self-selection. The seminal contribution of Borjas (1987) framed migration as an income-maximizing investment in human capital, predicting that migrants are not a random draw from their home population, but a self-selected group whose composition depends on the relative returns to skills in origin and destination countries. So far, the lack of administrative data on migrants' characteristics from origin country did not solve the question whether migrants are positively selected from origin country. Our contributions fills this gap in a context of low-income economy like Southern Italy.

Second, it extends the thriving literature on historical migration microdata pioneered by Abramitzky, Boustan, and Eriksson (2012) for the U.S. and advanced by Beck Knudsen (2024) for Scandinavia to a Southern European context. Mass emigration from Italy was distinct in its scale, regional character, and state management, offering a critical comparative case. This is the first paper to use administrative records from Ministry of Interior about passport applicants and also the first paper to study the selection and sorting criteria for a low-income area like the Southern Italy during the age of mass migration.

Third, it provides new empirical evidence on the interaction between migration policy, selection, and sorting. The seminal work by Grogger and Hanson (2011) formalized sorting as

an income-maximization process, where individuals choose destinations that offer the highest net return to their specific skill set. However, testing this hypothesis with historical data is hampered by the fragmented nature of traditional sources. Shipping manifests, for example, provide information only on migrants who completed a voyage to a specific port, creating a survival bias and obscuring the full range of intended destinations. Census data, while broader, rarely captures the pre-migration characteristics necessary to link individual traits to destination choice. Consequently, understanding of historical sorting patterns remains largely anecdotal or aggregated, missing the micro-level decision-making that connects individual agency to macroscopic flows. Administrative data from the country of origin of migrants help to identify the sorting criteria adopted by skill-heterogeneous migrants in the Age of Mass Migration according to two different migration policies. Indeed, we show that a more lenient migration policy expanded the emigrant pool spatially without degrading its average human capital but increased the pool of migrants towards countries with a greater labor demand in the construction and manufacturing industries.

The remainder of the paper is structured as follows. Section 2 reviews the theoretical and empirical literature on migrant selection and sorting, with a focus on the Italian case and historical data innovations. Section 3 describes our novel dataset, the archival digitization project, and our empirical methodology. Section 4 presents detailed descriptive statistics, constructing an "identikit" of the potential Italian emigrant. Section 5 presents our core econometric analysis of selection and sorting, including the impact of the 1901 policy shift. Section 6 discusses the implications of our findings for migration theory, historical narrative, and contemporary policy. Section 7 concludes.

2 Historical Background

The Italian Great Emigration, spanning roughly from the 1870s to the 1920s, was one of the largest mass migrations in modern history. An estimated 14 million Italians departed, with peaks occurring in the first decade of the 20th century. This exodus transformed the demographics of the several countries, in particular Americas (the United States, Argentina, and Brazil) and had profound, lasting impacts on the Italian economy, society, and state.

The role of the Italian state in managing this massive outflow is a critical and complex part of the story. Initially, the state viewed emigration with suspicion, as a drain on military manpower and a potential source of national shame (as emigrants often faced harsh conditions abroad). The 1888 law required passports and made emigration without official permission difficult, reflecting this restrictive stance. However, by the turn of the century, the state's perspective shifted. Emigrant remittances became a vital source of foreign currency, stabilizing the national currency and financing imports. Furthermore, emigration began to be seen as a solution to social unrest in the South. The 1901 Emigration Law, championed by Commissioner General of Emigration, Luigi Bodio, marked this liberal turn. It established a central emigration office, provided better information and some protection for emigrants, and effectively encouraged departure.

The seminal work by Faini and Venturini (1994) analyzed the determinants of emigration flows, emphasizing the role of wage differentials, demographic pressure (especially in the agrarian South), and the development of transnational transportation networks. They argued that while poverty was a push factor, it was not sufficient since the expectation of a higher income abroad and the means to achieve it (lowered transport costs, pre-existing networks) were crucial for the migration decision. This view was further refined and placed in a broader comparative context by Hatton and Williamson (1998). In their book, they developed a model

where emigration from Europe (including Italy) was driven by the widening wage gap between the Old and New Worlds, declining transport costs, and demographic transitions that created a "youth bulge" eager to seek opportunity abroad. For Italy, they highlighted the regional disparities: the industrializing North experienced a different, often smaller and more cyclical, emigration flow compared to the agrarian South, where the collapse of peasant agriculture and lack of industrialization created a persistent push. Our dataset, centered on Naples, captures the heart of this Southern emigration experience. The examination of individual characteristics of potential leavers allows us to study the key factors who influenced the decision of Italians to leave from different social contexts to different destination countries.

3 Data

3.1 Passport Applicants from National State Archive of Naples

The primary data is drawn from the *Questura* (Central Police Headquarters) folders of the passport applicants held at the State Archive of Naples (Archivio di Stato di Napoli). This collection comprises administrative records generated by the Naples police in the process of reviewing and issuing passports for international travel. As detailed in the project documents, the archive contains over 89,000 individual application folders, organized chronologically into major periods: 1881–1901 (over 20,000 folders), 1904–1909 (over 19,000), 1913–1917 (over 14,000), and 1926–1931 (over 26,000). The present analysis focuses on the first two periods which comprehend a major shift in Italian emigration.

We collect data from this archive since Naples was one of the main emigration port for the Italian emigration from 1888 onward. Naples served as the primary port of departure for transnational journeys from Southern Italy. Indeed, while the applications were processed in

Naples, applicants were not exclusively from the surround area. Individuals from across the Southern regions would often travel to Naples to submit their paperwork directly at the central Questura, rather than at their local municipal office, especially for long-distance overseas travel. As Perez (2021) notes, emigrants departing from Naples constituted approximately 10% of all Italian arrivals in Argentina and up to 50% of arrivals in the United States between 1880–1900. Therefore, while not a nationally representative sample of all Italian emigrants, this dataset provides a highly representative snapshot of the massive emigration stream from Southern Italy.

The 1888-1901 and 1904-1909 passport series contain more than 30,000 and 24,000 individual records. Each individual folder contains a standardized form for each applicant filled out by police officials. We collect several information from the form. First, we collect the following demographic and socioeconomic characteristics: gender, age, civic status, place of birth and residence at municipal level for the individuals residing out of Naples and the full address for those who live in Naples, literacy, self-declared occupation and family affiliation to other passport applicants. Then, we have collected the following new unique information on administrative and physical characteristics: criminal record, pending charges, military service status, height, eyes, hair, skin tone, name and surname of applicant, her/his mother and her/his father. Finally, we collect the following migration information: intended destination, reason for migration, date of passport application, a categorization of migrants for the bureaucratic process.

4 Descriptive Statistics

This section provides a descriptive evidence of potential leavers. Before showing the differences between passport holders and average population and the characteristics of passport holders among destinations, we provide a comprehensive identikit of passport applicants in the two periods of analysis, 1888-1901 and 1904-1909.

Figure 1 shows the number of passport applications for each from 1888 to 1909 with a gap of two years, 1902 and 1903. The number of sampled applications is stable from 1888 to 1896 around 2,000 observations per year. From 1897 to 1907 the number of applications grows monotonically. In particular, following the 1901, the number of applications increase exponentially reaching more than 5,000 sampled observations. Over the following year, the passport applications decreases at a level greater than before 1901. The analysis of the flows remarks the migration policy change occurred in 1901 and the amnesty granted from the Italian government to all the citizens who left Italy without holding a passport.

Figure 2 and Figure 3 show the age distribution of passport applicants by sex for the 1888-1901 and 1904-1909, respectively. The Subfigures (a), (b) and (c) show the total, male and female age distribution for both figures. Both figures show that individuals were more likely to migrate at early stage of their life, probably with their families, or at the adult age. Looking at the differences by gender, females were more likely to migrate during the first years of their life highlighting the gap between the total number of female and male emigrants during the adult age. Indeed, males were more likely to migrate at adult age when they finished the conscription and probably as temporary migrant since the share of female of the same age are less likely to migrate. Comparing the two figures, both men and women were more likely to apply for passport immediately after they reached the adult age in the 1904-1909 than in the previous period. The increase in the share of passport applicants in the period 1904-1909 is mirrored by a decrease in the share of individuals who apply for the passport later in their life, while the share of babies leaving with their families remain pretty much constant in both periods. An explanation would be that the migration rate of families grew constant, while the share of single emigrants aged 20-25 grew faster than those of older individuals.

Figure 4 shows the distribution of passport applicants among Italian province for 1888-

1901 (a) and 1904-1909 (b). In percentage, the chromatic differences are not much exacerbated between the two periods. The most pronounced difference is the change in the chromatic colors in the Southern regions. The provinces in Campania region become darker than the provinces of the other Southern regions.

Figure 5 shows the shares of passport applicants from each neighborhood in Naples in 1888-1901 (a) and 1904-1909 (b). The shares of passport applicants in the peripheral neighborhoods keep constant over time, while the shares change in the city center where the bottom neighborhoods become lighter and the inner ones become darker. The bottom neighborhoods are close to the coast and were richer, while the inner neighborhoods were poorer.

Figure 6 shows the share of passport applicants by occupation in 1888-1901 (a) and 1904-1909 (b). The comparison of the top occupations between periods show meaningful differences. While in the 1888-1901 period both men and women were massively employed in few specific occupation, housewife and tailor for females and cobbler, farmer, laborer and tailor for men, in the 1904-1909 males who apply for the passport are not concentrated within few occupations. Hairdresser, mechanic, nurse and waiter pop up as the new occupations of migrants probably going to the U.S.

5 Selection and Sorting

This section first provides some comparison between emigrants and stayers to study whether emigrants are positively selected from the population. Then, we will compare the emigrants among destinations to check whether skill-specific emigrants self-select in specific countries.

5.1 Selection

The 1888-1901 series contains 16,482 individual records for which occupation is available. Among those only around 10% declare of being agricultural workers, farmers or labourers, while the share of male agricultural workers was more than 50% in 1881 and 1901. In the same fashion, the 1904-1909 series shows that only around 6% of potential leavers were agricultural workers. Our data displays a strong positive selection of emigrants with respect to the average population in Italy. Our estimates might depend on the place where the data have been collected, the central police station of Naples. Therefore, comparing the estimates of shares of agricultural workers from our emigration data and the official census statistics for the district and the province of Naples, we find that the share of agricultural workers in the district is 4%, while it is 10% at provincial level for 1881. Therefore, provincial level statistics are more reliable since individuals from the municipality of Naples and surroundings included only 50% of individuals within the sample. Therefore, the share of agricultural workers in our dataset is equal to the share of agricultural workers of the province of Naples at least in 1881. Comparing the share of agricultural workers in our 1904-1909 sample with the population statistics in 1911, we observe that share of agricultural workers in the province of Naples was 20%, while the share of agricultural workers was still 10% in the 1904-1909 sample. Therefore, while in the 1888-1901 we do not observe any difference between the share of agricultural workers, in 1904-1909 we observe that share of emigrants working in agricultural industry was lower than the provincial mean. The analysis of workers in the agricultural industry highlights the null selection in the first emigration period and a positive selection in the second period.

Among workers in the manufacturing industry, individuals who self-declared workers in the textile industry were a consistent part of the total emigration in our sample. Indeed the 7% and the 15% of individuals defined themselves workers in the textile industry in 1888-1901

and 1904-1909 series, respectively. The share of workers in the overall manufacturing industry was the 27% in 1881 and 48% in 1911 in the province of Naples. The statistics show that the textile workers were a core part of the positive selection of emigrants in the 1904-1909 period. Summing the share of textile workers, a proto-industrialization job, and of agricultural workers, pre-industrialization job, we get that only 17% and 25% are occupations which are qualified at a low ranking of the occupational skills. Therefore, the rest of individuals are very likely to be more qualified than the individuals who declared the just mentioned occupations. This evidence highlights the positive selection of emigrants in both periods. As predictable, the selection was stronger in the more restrictive period than in the more lenient one.

5.2 Sorting

To understand how individuals sort across destinations, we present a set of tables showing the characteristics of individuals who declare their destination. First, Table 1 shows the shares of individuals within an occupation across countries for the 1881-1901 series. We narrow the table to only destinations with at least 200 individuals who apply for that destination and to occupations with at least 100 individuals declaring the same occupation. The main destination is the U.S. followed by Argentina, Brazil, North Africa (Algeria, Egypt and Tunisia) and European Countries (France, Greece and Russia). However, while passport applicants declaring a job in the textile and restaurant industries and barbers and hairdressers are more likely to declare the U.S. as their preferred destination, agricultural workers (farmers and laborers) and other manufacturer workers (carpenter, goldsmith, mason, stonemason) are more likely to travel towards Europe or North Africa. The U.S. were more attractive for workers with specific tasks in the textile industry and service sector, while other workers do not show strict preferences for specific destinations other than singers and stonemasons who travel in Europe. Male farmers are

more likely to go to South America with their wife, female farmers, than to the U.S. since the share of female farmers having as destination Argentina and Brazil are greater than the share of male farmers who are greater for the U.S.

Table 2 shows the shares of individuals within an occupation across countries for the 1904-1909 series. Also for this table, we narrow the table to only destinations with at least 200 individuals who apply for that destination and to occupations with at least 100 individuals declaring the same occupation. The vast majority of individuals go to the U.S., 73%, a statistics a way larger than the one observed in the previous series, 41%. The share of individuals going to the South America and the North Africa drops drastically, while the share of individuals going to European nations remain stable. The sorting of individuals with a specific occupation across destination does not appear in the data like in the previous series except for stonemasons who prefer Greece and landowners who split between France and the U.S. as all individuals decide to migrate to the U.S.. In particular, while agricultural workers were predominantly going to the South America, the same class of workers decide to go massively to the U.S. in the following period (96%). Finally, the new wave of passport applicants includes also a new team of professionals, the nurses.

Finally, Table 3 shows the mean of the age of passport applicants by destination. As before, we select only destination with at least 200 passport applicants in the two periods. Overall the average age of leavers decreases in the 1904-1909 period, the change in the average age depends mainly on the passport applicants to the US since the vast majority of them goes to the country in the North America. The average age of passport applicants to the U.S. decreases of more than 1.5 years. The decreases may depend on the 1901 migration Italian policy that fosters people to leave, the decrease in the transportation cost, the increase in the income per capita of Italians at beginning of the XX century and the increase in the average wages in the destination

country. The average age of passport applicants to other countries is greater except Switzerland. In particular, the average age jumps for Algeria and Brazil highlighting that passport applicants for those countries are less likely to move permanently.

6 Conclusion

This report analyses the migration preferences of passport applicants across several dimensions like age, occupation, place of birth and residence to delineate an identikit of potential leavers in a context of a developing economy like Italy in the late XIX century. We present descriptive evidences of the passport applicants in two different periods: the more stringent emigration policy period (1881-1901) and the more lenient one (1904-1909). The analysis allows us to compare different outflows looking at their destination preferences and their occupational status in the country of origin. We focus on the change in the type of selection and in the sorting of emigrants.

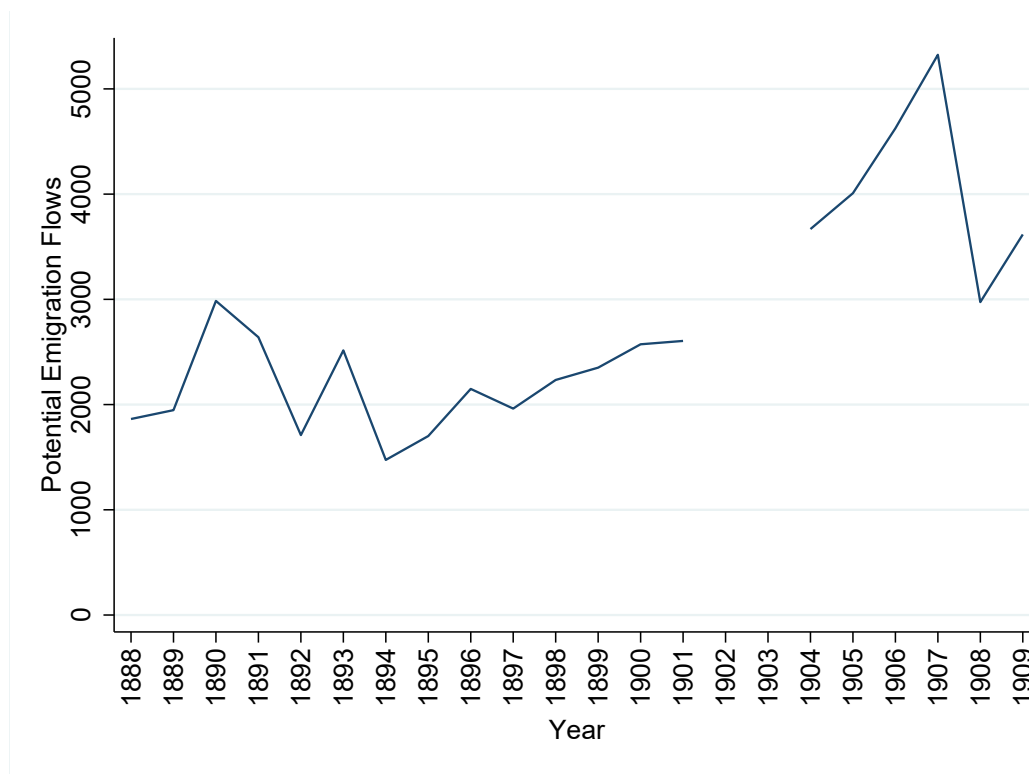
Our findings show that migrants are positively selected in both periods but we do find trade-offs among destinations when the insurance is not guaranteed for the richest country (the U.S. in our case). Passport applicants have a specific age range, 20-25, irrespective of their gender and are from the same areas of origin and residence in both periods irrespective of the income differentials. The meaningful differences stem from the occupation of potential leavers in the two periods. While the first series shows a predominance of farmers and tailors who go to different destinations, South America and North America respectively, in the second series passport applicants go massively to the U.S.. We do not find any heterogeneity in the destination between occupations for the individuals who apply for the passport between 1904-1909. Finally, our results suggest that a migration insurance provided by the government or the destination country pulls more young migrants since the age distribution between the two periods changes

based on the government emigration policy. During the period of a more stringent emigration policy, young emigrants prefer North America as much as South America since the cover of travel costs and the free land offset the greater earnings supplied in the U.S. Following the change to a more lenient emigration policy, the government provided to Italians job assistance once arrived in the U.S., supplying an insurance for Italians who decide to migrate massively to the U.S..

To conclude, this paper shows that emigration depends not only on future earnings and individuals skills but also on the incentives provided by the country of origin or destination to narrow the possible losses due to a bad sorting. Like in Brazil during the first wave of Italian emigrants at the end of the XIX century which government covered the travel cost and provided free land, the presence of new job office in the U.S. built by the Italian government at the beginning of XX century allowed many potential leavers to self-select in a country with higher earnings. Therefore, the role of assistance is a key driver of migration especially if the travel costs are larger than future potential earnings.

7 Figures and Tables

Figure 1: Flows of Passport Applicants from 1888 to 1909



Notes: The Figure shows the flows of potential leavers from 1888 to 1909. The figure lacks of data in 1902 and 1903 because of missing folders at the National State Archive of Naples for the two periods. We plot the figure using only the records who have at least one socio-demographic characteristics recorded other than the full name.

Figure 2: Age distribution of Passport Applicants by sex between 1888 and 1901

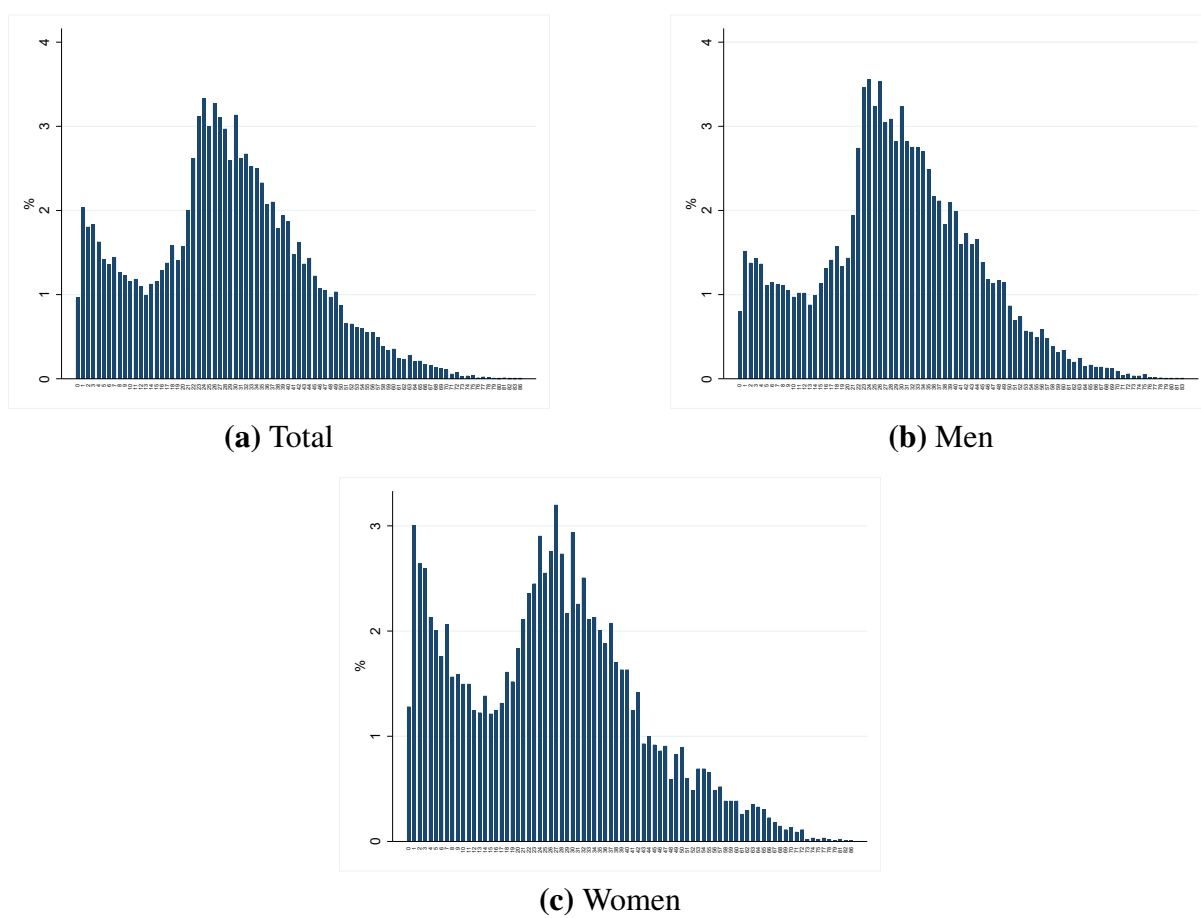


Figure 3: Age distribution of Passport Applicants by sex between 1904 and 1909

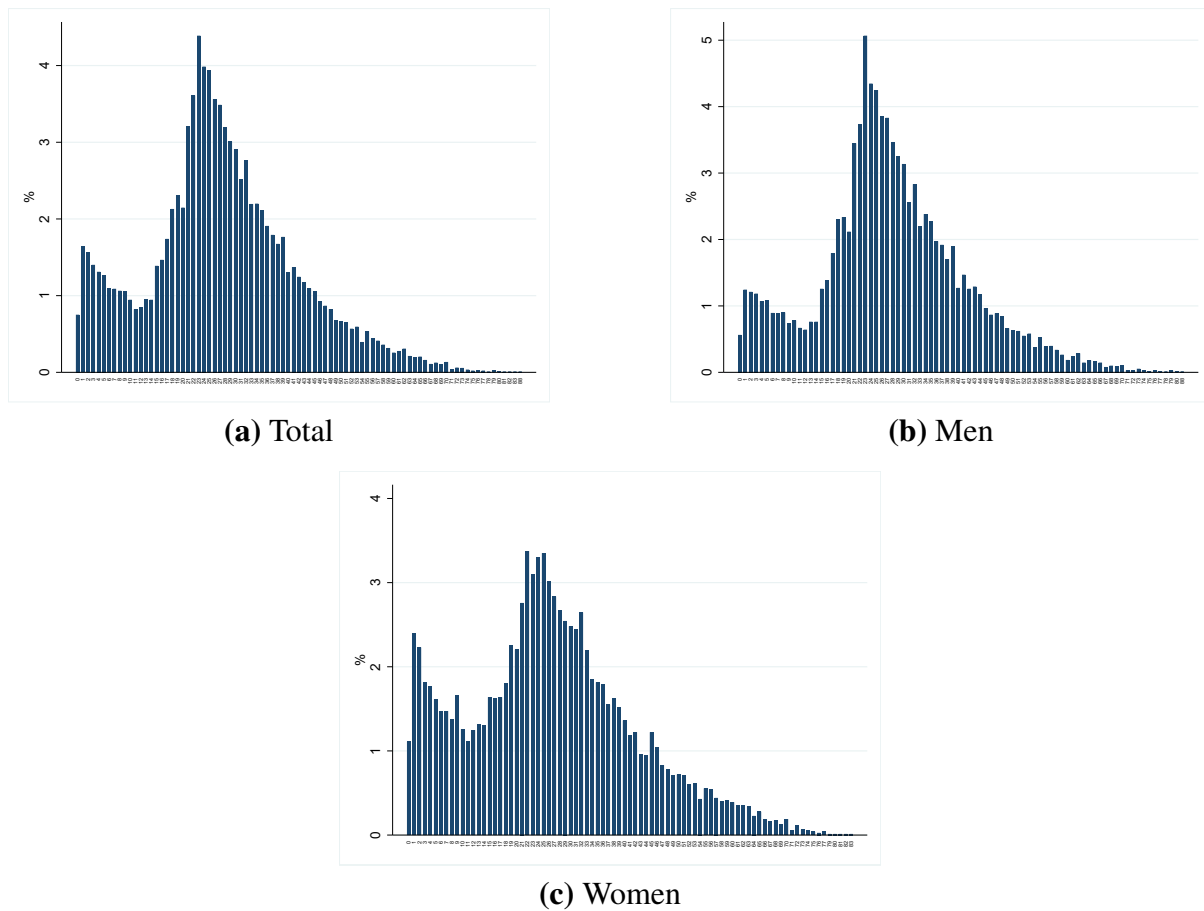
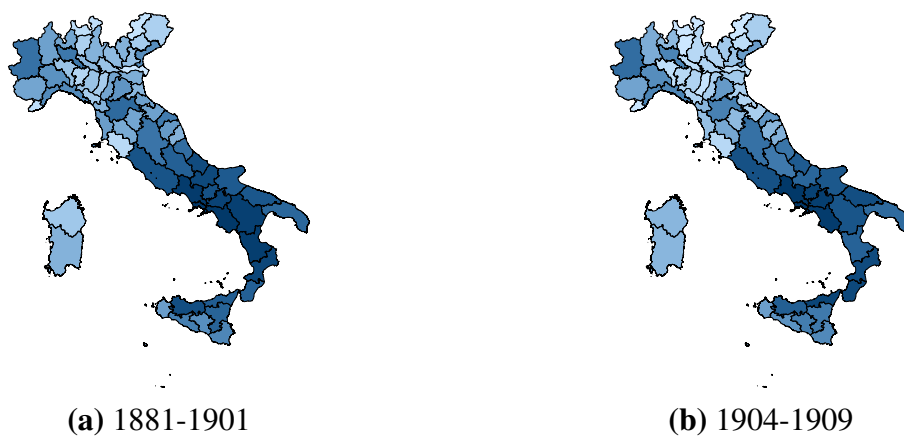
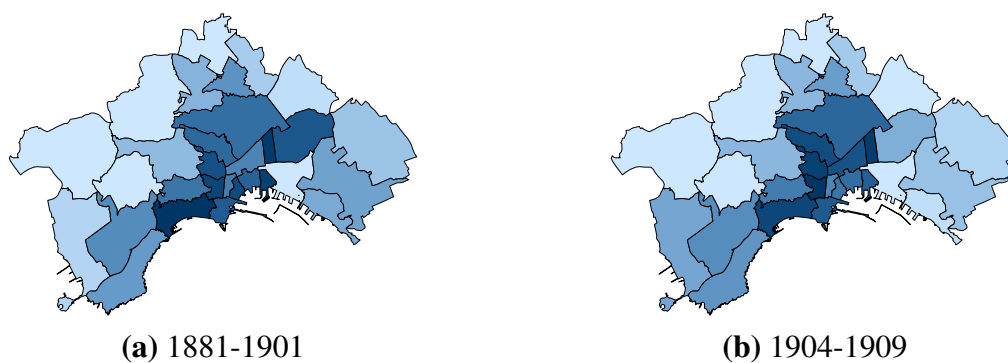


Figure 4: Number of Passport Applicants by Italian provinces



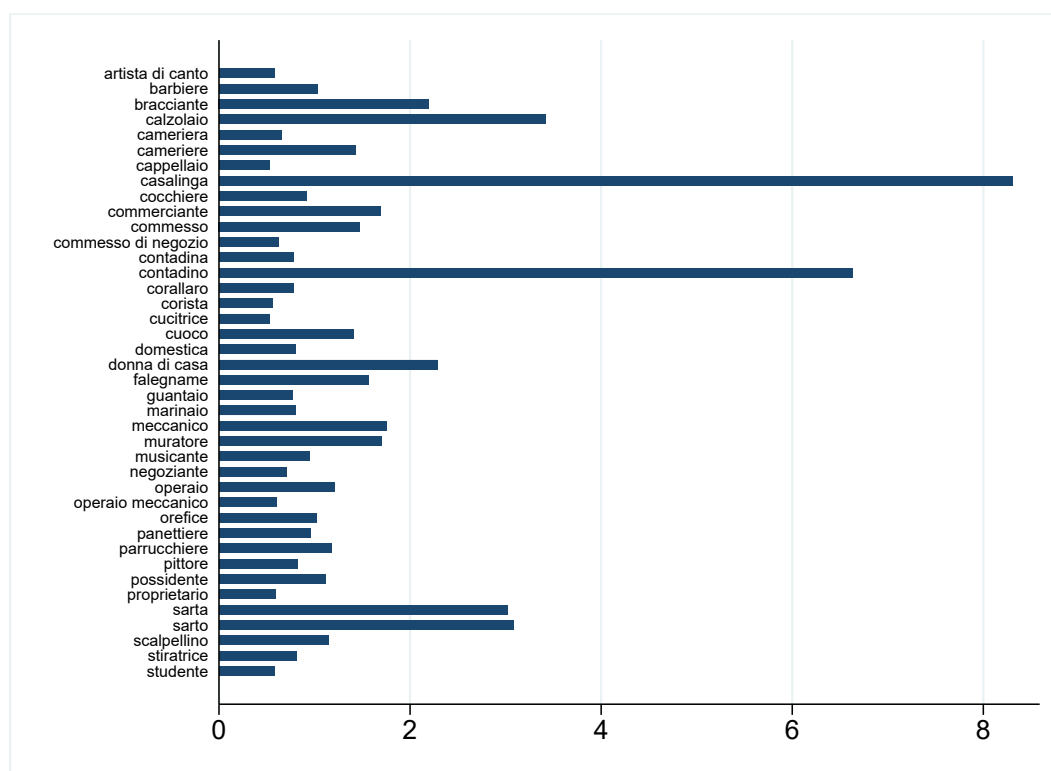
Notes: The Figure shows the distribution of immigrants across Italian provinces for two periods 1881-1901 and 1904-1909 in subfigures (a) and (b), respectively.

Figure 5: Share of Passport Applicants by Naples' Neighborhood

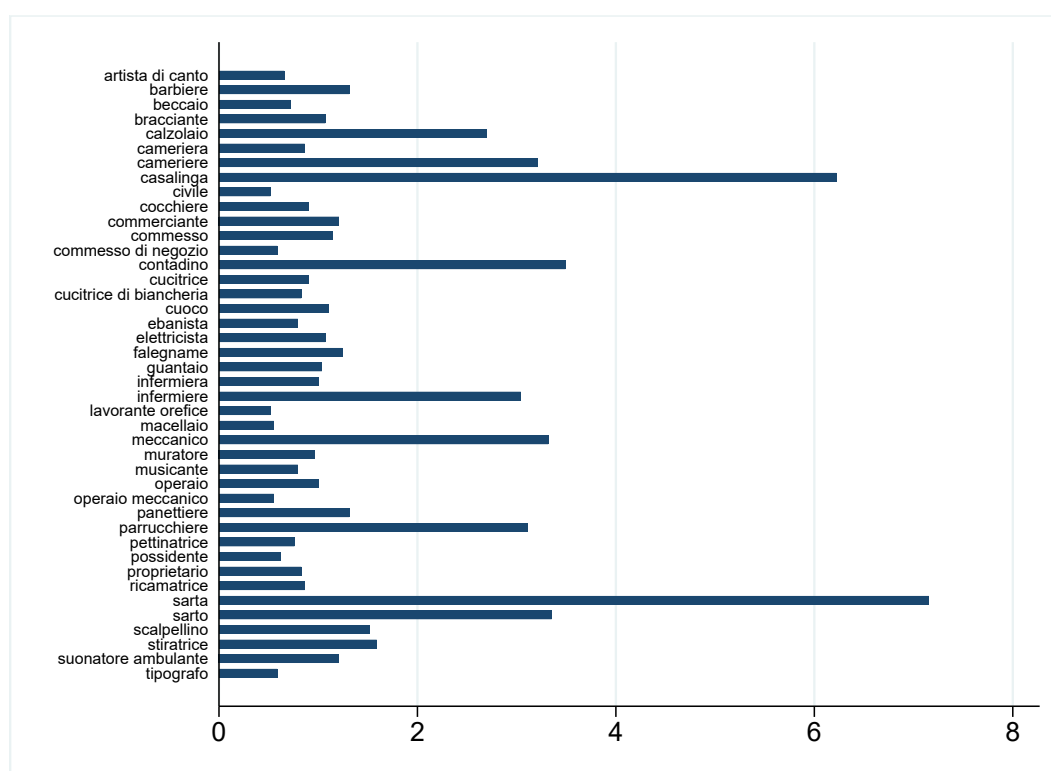


Notes: The Figure shows the distribution of immigrants across neighborhoods of residence in the subperiods. The subfigures (a) and (b) displays the emigration flows across Naples' neighborhoods between 1881-1901 and 1904-1909, respectively.

Figure 6: Share of Passport Applicants by Occupation



(a) 1881-1901



(b) 1904-1909

Notes: The Figure shows the distribution of immigrants across occupations in the subperiods. The subfigures (a) and (b) displays the most frequent occupations among potential leavers between 1881-1901 and 1904-1909, respectively. The individuals plotted account for more more than 60% of observations in both periods.

Table 1: Occupations' Distribution Across Countries Between 1881 and 1901

Occupation	Country										Total
	Algeria %	Argentina %	Brazil %	Egypt %	France %	Greece %	Russia %	Tunisia %	Uruguay %	USA %	
Baker	4	9	7	2	11	0	0	1	1	66	100
Barber	4	4	9	9	5	0	0	1	1	68	100
Carpenter	9	18	25	5	6	1	0	0	1	35	100
Clerk	1	19	14	13	10	3	0	1	3	35	100
Coachman	4	23	12	19	8	6	1	2	0	26	100
Cook	0	19	8	7	7	3	3	3	3	48	100
Coral Dealer	13	8	33	6	14	1	2	1	2	20	100
Female Farmer	1	13	63	4	3	0	0	0	2	15	100
Female Tailor	3	14	9	11	9	1	0	1	2	50	100
Glove Maker	0	6	5	3	11	3	0	0	0	72	100
Goldsmith	0	24	36	4	6	0	3	1	1	25	100
Hairdresser	2	14	5	10	10	1	1	1	0	57	100
Ironer	1	17	12	20	12	0	0	3	1	34	100
Laborer	8	10	27	3	10	3	0	1	2	37	100
Landowner	0	13	8	11	24	2	20	2	1	19	100
Maid	1	18	8	15	7	2	0	2	3	44	100
Male Farmer	1	9	50	0	3	0	0	0	1	36	100
Male Tailor	3	24	8	4	8	1	0	0	2	51	100
Mason	6	11	42	4	5	4	0	1	3	24	100
Mechanic	2	14	15	8	12	3	1	1	1	43	100
Merchant	1	20	13	11	11	3	4	0	2	36	100
Musician	0	6	8	13	14	17	6	1	3	31	100
Painter	3	21	13	5	13	12	2	0	0	32	100
Retailer	0	10	26	14	10	3	1	4	0	33	100
Sailor	18	24	5	2	18	1	0	2	0	30	100
Shoemaker	6	13	16	5	14	0	0	1	1	43	100
Shop Assistant	2	25	10	8	10	1	1	1	2	39	100
Singer	0	9	7	21	4	32	18	5	0	4	100
Stonemason	7	1	6	37	6	36	0	1	1	6	100
Waitress	0	20	9	17	8	1	3	0	1	40	100
Waiter	1	13	10	9	8	2	2	1	1	55	100
Worker	3	15	14	7	6	4	0	1	2	49	100
Total	4	14	19	8	9	3	1	1	1	41	100
N	284	1,098	1,490	617	675	213	83	72	107	3,269	7,908

Notes: Statistics are computed using only the occupations with at least 100 individuals who self-declare the same occupation and countries with at least 200 individuals who would like to go to the same destination.

Table 2: Occupations' Distribution Across Countries Between 1904 and 1909

Occupation	country									
	Algeria %	Argentina %	Brazil %	Egypt %	France %	Greece %	Russia %	Tunisia %	USA %	Total %
Baker	2	6	4	1	6	0	1	1	81	100
Barber	2	3	5	3	6	1	0	0	79	100
Busker	0	6	1	4	11	1	31	1	45	100
Carpenter	3	5	8	2	11	0	7	1	64	100
Coachman	2	12	8	11	14	5	0	1	47	100
Cook	0	10	3	2	8	2	1	0	75	100
Coral maker	6	16	10	0	3	1	3	2	58	100
Clerk	0	9	4	3	11	1	1	0	71	100
Electrician	1	9	3	3	7	0	1	0	76	100
Embroiderer	1	6	3	4	7	2	1	0	77	100
Farmer	1	1	1	0	2	0	0	0	96	100
Female Glove Maker	0	6	2	2	1	2	0	0	87	100
Female Hairdresser	1	4	5	1	12	2	0	0	76	100
Female Nurse	0	3	2	0	1	0	0	0	95	100
Female Tailor	1	6	4	2	7	0	0	0	78	100
Ironer	2	7	7	3	11	2	1	0	66	100
Laborer	4	4	4	2	24	5	0	1	58	100
Landowner	0	2	3	9	35	2	11	0	39	100
Linen Seamstress	1	6	2	2	5	0	1	1	83	100
Male Glove Maker	0	6	1	0	9	2	0	1	82	100
Male Hairdresser	0	5	4	1	7	0	0	1	82	100
Male Nurse	0	3	4	1	2	1	0	0	91	100
Male Tailor	1	5	4	1	13	1	1	0	74	100
Mason	3	4	10	1	8	1	1	0	72	100
Mechanic	1	7	5	2	9	1	1	3	71	100
Mechanic Worker	0	6	6	3	7	0	0	1	77	100
Merchant	1	9	9	11	14	1	4	3	49	100
Musician	0	15	0	8	13	0	19	0	44	100
Office Worker	0	13	2	9	9	2	0	0	65	100
Painter	1	14	3	6	13	0	0	0	64	100
Sailor	19	15	0	0	2	0	1	5	58	100
Seamstress	5	4	6	2	8	2	0	1	72	100
Shoemaker	5	4	4	1	21	1	0	1	63	100
Stonemason	3	1	4	2	2	74	0	1	14	100
Typographer	3	2	6	2	8	1	0	0	79	100
Waitress	1	7	4	5	5	1	0	0	78	100
Waiter	1	5	3	1	9	0	0	0	81	100
Worker	4	4	7	2	15	7	0	1	60	100
Total	2	6	4	3	9	3	1	1	73	100
N	163	545	403	252	840	277	115	79	7,184	9,858

Notes: Statistics are computed using only the occupations with at least 100 individuals who self-declare the same occupation and countries with at least 200 individuals who would like to go to the same destination.

Table 3: Average Age of Passport Applicants in 1881-1901 and 1904-1909

Country	1881-1901	1904-1909
Algeria	27.5	31.1
Argentina	28.8	28.3
Brazil	26.4	29.5
Egypt	31.0	30.0
France	29.1	28.8
Germany	28.7	28.9
Greece	34.2	34.9
Russia	31.8	30.2
Switzerland	30.7	27.4
Tunisia	30.3	30.4
Turkey	30.2	37.2
US	27.7	25.9
Total	28.3	27.0

Notes: Age means are computed using only destination countries with at least 100 individuals.

Logged Out: How COVID-19 School Closures Deepened Educational Poverty in Italy

Emma Sarno

Department of Human and Social Sciences,
University of Naples “L’Orientale”
`esarno@unior.it`

Abstract

This paper examines how the digital divide intensified educational inequalities in Italy after the unplanned transition to remote teaching during the COVID-19 pandemic. It distinguishes three phases: the pre-pandemic baseline of 2018–2019, the pandemic shock of 2020–2021, and post-pandemic evidence on educational outcomes. Combining micro- and macro-level data on socioeconomic status, territorial disparities, digital access, and student achievement, the study shows that remote teaching was not a neutral emergency response. Instead, it magnified pre-existing inequalities, especially along Italy’s North–South divide. Students from disadvantaged households and Southern regions faced greater barriers to devices, connectivity, and adequate learning environments. The findings suggest that digital exclusion operated as a structural mechanism through which social and territorial inequalities were reproduced, translating unequal access to technology into unequal learning opportunities and, ultimately, into deeper forms of educational poverty.

Keywords: digital infrastructure; distance learning; students’ achievement; educational poverty

JEL codes: I24, I18, I3.

1 Introduction

This paper investigates how digital infrastructure, access to digital devices, digital competences, domestic learning conditions, and school-level organization shaped educational opportunities in Italy during the COVID-19 pandemic. The pandemic produced an abrupt acceleration in the digitalization of teaching and learning activities, as schools were forced to reorganize educational provision in response to social distancing measures. Yet this acceleration did not take place in a neutral environment. It occurred within an education system already marked by persistent territorial, socio-economic, and performance gaps, and therefore provides a critical context in which to assess whether digitalization operated as an equalizing force or as a mechanism for reproducing pre-existing inequalities.

The shift was sudden. In March 2020, in-person educational activities were suspended nationwide, and schools and universities moved from traditional face-to-face instruction to remote teaching under emergency conditions. During the remainder of the 2019–2020 school year, remote teaching became the main channel through which educational continuity was maintained. In the 2020–2021 school year, online instruction continued to play a major role, although it was combined with periods of in-person attendance and with blended arrangements based on student rotation, subgrouping, school-building capacity, and local epidemiological conditions. While the first phase of school closures was imposed uniformly across the national territory, during the second wave decisions on partial or full suspension of in-person teaching became increasingly differentiated by region, school level, and local health conditions. Remote teaching was therefore far from homogeneous. Evidence from Bovini and De Philippis (2021), based on a Bank of Italy survey, shows that between March 2020 and February 2021 the organization of distance learning varied substantially across school levels and regions. In primary and lower-secondary schools, almost half of distance-learning hours were delivered asynchronously, meaning that teachers and students were not simultaneously connected to the same digital platform. In upper-secondary schools, by contrast, remote teaching was more frequently delivered synchronously. Moreover, the regions and school levels where in-person teaching was reduced the most tended to compensate by increasing the share of synchronous online lessons. This was particularly evident in Southern regions and, above all, in Campania, where regional ordinances suspended face-to-face teaching for longer than elsewhere.

In principle, digital education can offer important opportunities for learning. Digital pedagogies may support new forms of interaction, extend access to educational resources, strengthen communication between teachers and students, and enable the use of online platforms, videoconferencing tools, and digital learning environments. However, digital education does not operate in a social vacuum. Its effectiveness depends on a set of

preconditions that are unequally distributed across households, schools, and territories: reliable connectivity, appropriate devices, adequate digital skills, suitable domestic spaces, and family support where needed. The central question is therefore not whether digital technologies can improve education in abstract terms, but under what conditions they do so, and for whom.

This question is particularly relevant in the Italian case. Italy entered the pandemic with long-standing inequalities in educational performance, socio-economic background, school trajectories, and territorial development. The North–South divide, in particular, continues to structure both students’ learning outcomes and their access to material and technological resources. In such a context, the transition to remote teaching exposed a crucial tension within the digitalization of education: the same technologies that may expand learning opportunities for some students may also amplify exclusion for those who lack the material, cognitive, and domestic conditions required to benefit from them.

A growing international literature has assessed the educational consequences of COVID-19 school closures and remote teaching. Existing evidence shows that school closures reduced instructional time, generated learning losses, and widened pre-existing educational inequalities, especially for students from disadvantaged socio-economic backgrounds. Engzell, Frey, and Verhagen (2021), using evidence from the Netherlands, show that learning losses were substantial even in a country characterized by a relatively short period of school closures and high levels of digital preparedness, with stronger effects among pupils from less-educated households. Global simulations by Azevedo et al. (2021) similarly suggest that the disruption reduced schooling and learning outcomes, increased dropout risks, and may have long-term consequences for future earnings and human capital accumulation. From the perspective of the economics of education, Hanushek and Woessmann (2020) also stress that even temporary learning losses can have persistent effects on individual lifetime income and aggregate economic performance.

A second strand of the literature has emphasized that the emergency shift to online instruction should not be confused with well-designed online learning. Hodges et al. (2020) define this crisis-driven response as “emergency remote teaching”, a temporary solution adopted under conditions of urgency rather than as the result of deliberate pedagogical planning. This distinction is crucial for interpreting the Italian experience. Distance learning was introduced under severe time constraints, with heterogeneous levels of digital infrastructure, teacher preparedness, organizational capacity, and family support. Its effectiveness therefore depended not only on access to platforms and devices, but also on the extent to which technologies could be embedded in coherent pedagogical practices, redesigned learning environments, and adequate forms of teacher-student interaction.

Recent empirical research on the Italian case points to a similar ambivalence. The

emergency shift to remote education produced uneven effects across schools, students, and households, depending on infrastructure, teachers' digital and pedagogical competences, students' autonomy, household resources, and territorial conditions. Lo Presti and Dentale (2025) interpret the pandemic as a large-scale collective experiment in digital teaching, showing how schools were forced to reorganize access to education, curricular practices, assessment, social relations, and school-family relationships under unequal conditions. Contributions in the same volume specify these mechanisms from different perspectives: Martire, Palmieri, Mastrolia, and Seller (2025) emphasize technological endowments, broadband stability, platform availability, and teacher training; Lo Presti, Capozza, Salvi, and Cavallo (2025) show that digital tools did not automatically generate pedagogical innovation; Fasanella, Dentale, Capozza, and Rimoli (2025) point to mixed effects on competence-based learning; and Lo Presti, Faggiano, and Dentale (2025) highlight the relational and socio-emotional consequences of remote education. The present analysis moves in the same direction, interpreting these dimensions as part of a cumulative process through which digital exclusion contributed to educational inequality and educational poverty.

This article also builds on previous research (Sarno, 2022), which provided an initial assessment of the Italian experience with remote teaching during the emergency phase of the pandemic, focusing on the period up to 2021. The present paper extends that earlier analysis in both temporal and analytical terms. Temporally, it connects the immediate disruption caused by remote schooling to the pre-pandemic structural baseline of 2018–2019, and to post-pandemic evidence on educational outcomes and educational poverty. Analytically, it shifts the focus from the emergency management of remote teaching to the mechanisms through which digital exclusion translated social and territorial inequalities into unequal learning opportunities.

The paper therefore adopts a multi-layered research design structured around three analytical moments. First, it reconstructs the pre-pandemic baseline by examining digital infrastructure, device availability, housing constraints, and educational performance before the outbreak of COVID-19. Second, it analyzes the contextual shock of the 2020–2021 school closures, focusing on students' participation in remote teaching and on the unequal conditions under which distance learning took place. Third, it examines post-pandemic evidence on the longer-term consequences of the crisis, with particular attention to learning outcomes and educational poverty.

The argument developed in the paper is that remote teaching did not function as a neutral emergency substitute for in-person schooling. Rather, it interacted with pre-existing inequalities and, in some cases, intensified them. Connectivity, device availability, housing conditions, family support, and school-track stratification did not operate as isolated

variables; they combined cumulatively, producing a self-reinforcing mechanism of exclusion. In this sense, the Italian case can be read through the lens of cumulative causality, or a “Matthew Effect”, whereby students and territories already endowed with stronger resources were better positioned to maintain educational continuity, while disadvantaged students faced a higher risk of disconnection, learning loss, disengagement, and educational deprivation.

The evidence points to a familiar but deeply consequential image of Italy: a country divided along territorial and socio-economic lines, where family background and place of residence continue to shape both educational performance and access to innovative technologies. In a context where remote teaching became an essential component of educational provision, the digital divide did not merely represent a temporary technical obstacle. It became a structural channel through which existing inequalities were reproduced and translated into unequal learning opportunities. From this perspective, digital exclusion contributed not only to lower achievement, but also to deeper forms of educational poverty, understood as the deprivation of the opportunities, resources, and capabilities needed to learn, participate, and fully develop one’s potential.

The remainder of the paper is organized as follows. Section 2 examines the material preconditions of remote teaching by bringing together digital infrastructure, connectivity, device availability, housing conditions, study space, and family support. Section 3 shifts the focus to students, linking pre-pandemic achievement and digital skills to participation in remote teaching, perceived learning effectiveness, school-track differences, and post-pandemic indicators of achievement, dropout, and educational poverty. Section 4 concludes by drawing together the mechanisms examined in the two analytical sections and by reflecting on the relationship between digital divide, educational inequality, and educational poverty in post-pandemic Italy.

2 Digital and domestic conditions before and during the pandemic

This section reconstructs the material preconditions that shaped the feasibility of remote teaching in Italy. It first examines digital infrastructure and digital readiness before the pandemic, then considers how territorial and socio-economic inequalities affected connectivity and access to devices, and finally turns to the domestic learning environments in which remote schooling actually took place.

2.1 Digital infrastructure and digital readiness

To trace the structural origin of the educational poverty spiral, it is necessary to establish the pre-pandemic baseline of Italy’s digital ecosystem. By evaluating the structural and human capital indicators immediately preceding the crisis, we can map the systemic vulnerabilities that later conditioned and restricted the deployment of emergency remote teaching.

Since 2014, the European Commission has monitored Member States’ digital performance through the Digital Economy and Society Index (DESI). DESI¹ is a composite indicator based on several measures of digital competitiveness. It covers five main dimensions: i) *connectivity*, to summarize the deployment and quality of broadband infrastructure; ii) *human capital*, to capture internet use and basic and advanced digital skills; iii) *use of internet services*, to measure the range of online activities undertaken by citizens; iv) *integration of digital technology*, to measure business digitalization and e-commerce; v) *digital public services*, to capture the digitalization of public services and e-government.

DESI data for 2020 show Finland as the digital leader, with a score of 72.3, followed by Sweden and Denmark. Italy, with a score of 43.6, ranks 25th out of 28 EU Member States, ahead only of Romania, Greece and Bulgaria. For the two DESI dimensions most relevant to education, Italy presents a mixed picture. Its connectivity score is 50, only slightly below the EU average, placing the country 17th among EU Member States. Conversely, Italy demonstrates the weakest performance within Europe regarding human capital. Specifically, only 42% of individuals aged 16–74 possess at least basic digital skills, compared to the European Union average of 58%. Furthermore, a mere 22% exhibit above-basic digital proficiency, in stark contrast to the broader EU average of 33%. Crucially, this demographic bracket encompasses the core actors of the educational ecosystem, both students and educators, highlighting the significant risk that those tasked with navigating the crisis were inadequately prepared for the sudden transition to remote learning on one side and teaching on the other. The number of ICT specialists and ICT graduates is also well below the EU average.

¹From 2014 to 2022, DESI provided a composite index summarizing countries’ performance as described in the above text. Later, a methodological change occurred in 2023 following the implementation of the Digital Decade Policy Programme 2030. As a result, DESI ceased to be published as a separate composite index and was integrated into the annual State of the Digital Decade Report. The focus shifted from producing a single summary score toward monitoring progress against specific Digital Decade targets and key performance indicators (KPIs). Consequently, DESI values published up to 2022 are not directly comparable with the indicators reported from 2023 onwards. However, this discontinuity does not affect the period examined here as it stops before then.

2.2 Connectivity and digital devices

The disappointing picture provided by the DESI index conceals substantial differences across Italian regions, exposing the geographical and educational fractures that fed the cumulative downward spiral during the school closures. The ISTAT multipurpose survey, *Aspects of Daily Life*, includes a section on *Citizens and ICT* ² that reports the share of households with a broadband connection. Data for 2018 and 2019 indicate an overall improvement in connectivity, with the national share of households connected to broadband rising from 72.7% to 74.0%. The increase was stronger in the Centre, but differences between macro-regions also widened slightly, confirming that structural progress was inherently unequal (Table 1).

Table 1: Broadband connectivity of Italian households (in percentage)

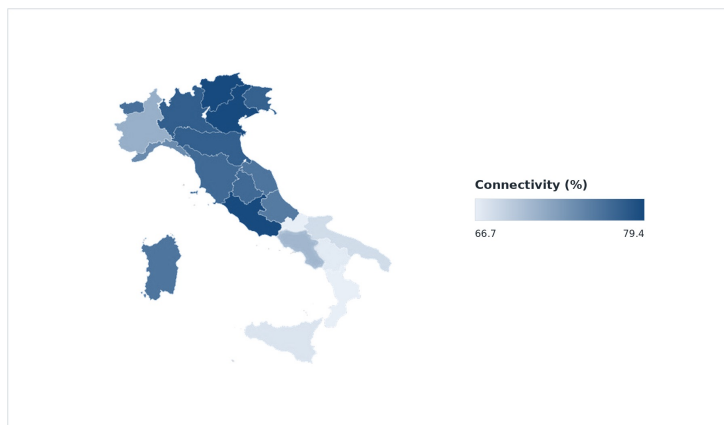
Macro-Region	2018	2019
Italy (μ)	72.7	74.0
North	75.5	76.5
Centre	75.1	77.0
South	68.6	69.5
Islands	71.0	71.8

Source: ISTAT.

Consistent with these data, the connectivity map in Figure 1 highlights a clear regional divide in broadband diffusion. All southern regions, with the exception of Abruzzo, lag behind the northern and central regions. The gap is 12.7 percentage points between Trentino, the region with the highest share of households connected to broadband, and Molise, the region with the lowest share of households connected to broadband. The share of households with a broadband connection also varies by the average education level and age composition of household members, illustrating how class boundaries reinforce technological isolation. Only 5.9% of households in which at least one member holds a university degree lack a broadband connection, whereas the share rises to 53.9% among households whose members have only a lower secondary qualification. Similarly, 66% of households composed exclusively of older people have no connection, while this share falls to 4.9% when at least one minor is present.

²Istat (2019) Report: Cittadini e ICT, <https://www.istat.it/it/files//2019/12/Cittadini-e-ICT-2019.pdf>

Figure 1: Regional levels of connectivity



Source: Author's elaboration based on ISTAT data (2019).

A second key condition for assessing the potential effectiveness of remote teaching is the availability of computers and tablets. Ideally, each student should have access to a personal device. According to ISTAT ³, in 2018-2019, 33.8% of Italian households had neither a computer nor a tablet; in the South, this share reached a critical 41%. Among households with at least one minor, the figure fell to 14.3%. More importantly, 12.3% of children aged 6-17 had no computer or tablet at home, corresponding to approximately 850,000 children and adolescents. In the South, the share was almost 20%, or about 470,000 students - more than half of all Italian students without a device. Only 22.2% of households had one PC or tablet for each member, and the share was just 14.1% in the South. Personal devices were available to only 6.1% of children living in households with at least one computer per household member.

In summary, although around 95% of households with minors were connected, distance learning was severely constrained by an insufficient number of PCs and tablets relative to the number of household members. A substantial share of students therefore had to share hardware with other family members or, worse, attend online lessons using only a smartphone. Significant inequalities between the North and the South are evident both in connectivity and in access to ICT devices, forming a rigid baseline of material exclusion.

2.3 Housing overcrowding, study space, and family support

The spatial bottleneck and the domestic environment constitute critical intermediate vectors that accelerated educational marginalization. When emergency remote teaching abruptly shifted the educational process from physical classrooms to private households, the

³Istat (2020) Report: Spazi in casa e disponibilità di computer per bambini e ragazzi, <https://www.istat.it/it/files//2020/04/Spazi-casa-disponibilita-computer-ragazzi.pdf>

material and relational resources available within domestic spaces immediately translated into asymmetric learning barriers. The transition from face-to-face to remote teaching changed the physical space in which learning occurred. On the one hand, distance learning allows students to attend school or university from any location, thereby reducing mobility costs. On the other hand, this apparent spatial freedom is limited by the quality of the home environment. Students do not always have suitable conditions for learning because of overcrowded housing or insufficient support from adult household members. This spatial bottleneck was a documented global challenge during the pandemic; as highlighted by the OECD (2020), housing quality quickly transformed from a private welfare indicator into a primary determinant of educational equity.

Eurostat defines overcrowding according to the bedroom standard, which assigns more demanding space requirements as children grow older and varies them by gender composition⁴. In 2019, on the eve of the pandemic, 28.3% of the Italian population lived in overcrowded housing. The situation was worse for minors, particularly those in low-income households. Table 2 distinguishes between poor households (defined as those with income below 60% of the median equivalized income) and non-poor households. More than half of poor households with minors experienced overcrowding⁵. Among non-poor households, the overcrowding rate was 34.1% for children under 6 and increased to 44.9% for children aged 12-17, reflecting the stricter bedroom standard applied to this age group.

Table 2: Overcrowding rates among minors by age and economic condition

Minors	Above 60% of the median equivalized income	Below 60% of the median equivalized income
0-6 years	34.1	46.2
6-12 years	36.5	51.6
12-17 years	44.9	53.0

Source: EU-SILC, 2019.

⁴Precisely, Eurostat definition of overcrowding states that an overcrowded household is one which has fewer rooms than the sum of: one room for the household as a whole; one room per couple in the household; one room for each single person aged 18 or more; one room per pair of single people of the same gender between 12 and 17 years of age; one room for each single person of different gender between 12 and 17 years of age; one room per pair of children under 12 years of age.

⁵Source: European Union Statistics on Income and Living Conditions (EU-SILC), Overcrowding rate by age, sex and poverty status (ILC_LVHO05A). Although overcrowding has declined somewhat in recent years, with the average overcrowding rate falling to 23.9% in 2024, Italy continues to display comparatively high levels of housing crowding relative to other major European countries, especially among low-income households.

A further obstacle to effective distance learning is the lack of family support for early-grade primary pupils and for students with disabilities. This vulnerability aligns closely with international literature tracking the unequal toll of remote learning across Europe. Evidence from the Netherlands—a country with near-universal broadband access—demonstrated that even short periods of school closures resulted in substantial learning losses that were up to 60% larger among students from less-educated homes (Engzell et al., 2021). Similarly, research in the UK showed that parental resources and available time directly conditioned the quantity and quality of home-learning hours (Andrew et al., 2020).

In the Italian context, these dynamics were compounded by structural factors. Under ordinary circumstances, the education system is expected to provide assistance through classroom teachers or specialized support educators (*insegnanti di sostegno*). Distance learning increased childcare responsibilities, especially for women (Bovini and De Philippis, 2021), thereby intensifying unpaid work within the household. However, family members do not always have the time, skills, or resources required to provide effective support, causing initial class disadvantages to multiply. An *ad hoc* ISTAT survey on the inclusion of pupils with disabilities⁶ shows that almost 300,000 students with disabilities attend Italian schools, corresponding to 3.5% of enrolled pupils. Between April and June 2020, more than 23% of them — about 70,000 students — did not participate in remote lessons; in the South, this share reached a critical 29%. The survey identifies several reasons that illustrate how different forms of deprivation reinforce each other: the severity of the condition (27%), difficulties in cooperation with family members (20%), socio-economic hardship (17%), difficulties in adapting the Individual Education Plan to distance learning (6%), lack of technological tools (6%), and a lack of specific teaching aids (3%).

3 Students’ experiences and educational outcomes

This second analytical section shifts from material preconditions to students’ experiences and outcomes. It first reconstructs students’ achievement and digital skills before the pandemic, then examines participation in remote teaching during the first lockdown, and finally connects perceived learning effectiveness and school-track differences to post-pandemic indicators of dropout and educational poverty.

A comprehensive assessment of the impact of remote teaching on students in Italian schools is constrained by the lack of representative data for the entire education system. This prevents a fully accurate evaluation that accounts for the functional and geographical complexity of Italian schooling. Nevertheless, the available evidence makes it possible

⁶ISTAT (2020), Report: L’inclusione scolastica degli alunni con disabilità - A.A. 2019-2020, <https://www.istat.it/it/archivio/251409>

to reconstruct the learning conditions in place when the pandemic emergency and the spring 2020 school lockdown occurred, and to identify the challenges that distance learning introduced.

3.1 Evidence from OECD PISA and INVALSI on student performance

This subsection maps the baseline cognitive and functional literacy constraints of Italian students prior to the crisis. These pre-existing educational divides directly restricted the students' cognitive readiness to transition from guided physical instruction to autonomous digital knowledge construction.

To evaluate education systems, the OECD launched the Programme for International Student Assessment (PISA) in 2000. PISA measures the performance of 15-year-old students in mathematics, science, and reading. Repeated every three years, the survey provides internationally comparable data designed to support improvements in education policies and outcomes. In Italy, additional evidence is collected by INVALSI (*Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione*), whose specific aim is to assess students' competences across all levels and grades of the Italian school system.

In 2018, Italy scored below the OECD PISA average in reading and science and around the OECD average in mathematics⁷. After 2012, average scores declined in reading and science, while mathematics remained broadly stable and above the levels observed in 2003 and 2006. The decline in reading performance was concentrated among girls, while boys' performance remained stable. The decline in science was particularly marked among high-achieving students and was similar across gender.

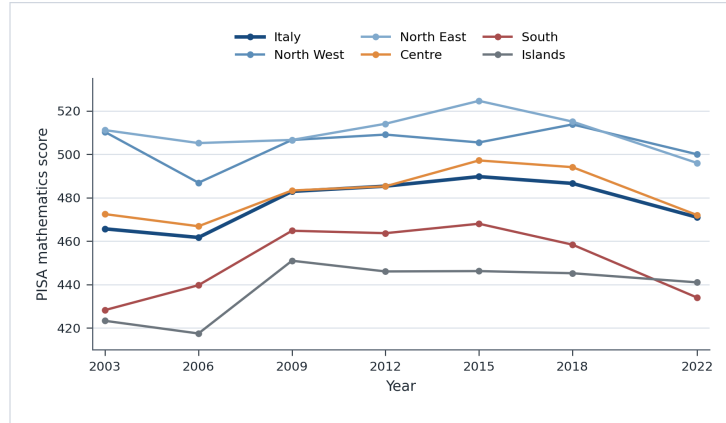
The career expectations of the highest-achieving 15-year-old students reveal strong gender stereotypes. Among high performers in mathematics or science, about one in four boys in Italy expected to work as an engineer or science professional by age 30, compared with only one in eight girls. Only 7% of boys, and almost no girls, expected to work in ICT-related occupations, providing further evidence of Italy's limited overall orientation toward digital fields.

Students' performance is strongly associated with family socioeconomic background, as lower scores are correlated with low household income and low parental education. Weaker educational profiles are also often observed among students with a migration background, pointing to integration difficulties that may contribute to economic and social

⁷OCSE PISA 2018, https://public.tableau.com/app/profile/INVALSI/viz/OCSEPISA2018_15765811689360/Copertina

exclusion. Figure 2 reports Italian mathematics scores since the beginning of OECD PISA and shows pronounced geographical differences, a pattern that also characterizes reading and science. Students in the North systematically score above the national average, those in the South and Islands systematically score below it, while students in the Centre occupy an intermediate position.

Figure 2: Student performance by Italian macro-region, 2003-2022



Source: Author's elaboration based on OECD PISA and INVALSI data.

The 2022 data further reinforce this interpretation by showing a twofold post-pandemic effect: a general deterioration in students' achievement and a further differentiation among macro-areas. The decline in performance appears to be a national phenomenon, but it is more pronounced in the South, where pre-existing educational disadvantages were already stronger. Thus, the pandemic did not affect all students and territories in the same way; rather, it amplified existing regional inequalities and made the territorial divide in educational outcomes even more evident.

Crucially, these regional and socio-economic disparities in academic domains established a fragile baseline. When educational delivery shifted to the home, existing cognitive inequalities directly merged with technological deficits, leaving specific demographics uniquely vulnerable. Consequently, the traditional geographical fracture between the North and the South served as a predictive map for how severe the digital exclusion would be once remote teaching commenced.

Drawing on Potter and McDougall's (2017) concept of digital literacy, the subsequent PISA insights (OECD, 2021) state that literacy in the twentieth century concerned the extraction and processing of pre-coded and usually curated information, whereas in the twenty-first century it concerns the construction and validation of knowledge. For this reason, OECD PISA recognized the importance of assessing students' digital skills. Although the 2021 survey cycle was postponed to 2022 to allow sufficient time for implementation,

preliminary frameworks highlight a close relationship between reading performance and digital skills. They also indicate that digital divides mirror existing economic gaps and may amplify the disadvantages of students from less affluent backgrounds, thereby widening learning inequalities.

This weakness is corroborated by international large-scale assessments specifically targeting digital literacy. According to the 2018 International Computer and Information Literacy Study (ICILS), which measures the skills of 13-year-old pupils through computer-based tests, Italian students scored an average of 461 points. This is considerably lower than their peers in other European countries, such as Denmark (553), Finland (531), and Germany (518). Furthermore, 24% of Italian students failed to reach the minimum threshold of basic digital skills, revealing a substantial gap compared to Denmark (3%), Finland (8%), and France (10%) (Fraillon et al., 2020).

The fact that nearly a quarter of Italian 13-year-olds lacked foundational digital literacy prior to the crisis underscores that the digital divide was not only an issue of hardware availability, but also of functional capability. In a system where one in four students could not navigate a basic computer-based environment autonomously, the sudden, absolute reliance on remote teaching meant that a substantial segment of the student population was structurally and cognitively locked out of effective learning from day one. This 24% baseline deficit represents a critical mass of students who were functionally unprepared for the shift from passive information consumption to active digital knowledge construction.

These macro-level structural gaps align closely with localized findings. Save the Children (2020), in a pilot study on digital educational poverty among 13-year-old students in selected Italian provinces⁸, showed that about 20% of participating children were unable to perform simple operations using ICT tools. For instance, 11% could not share a screen during a Zoom call, and 29.3% were unable to download a document shared by a teacher on the school platform. Overall, 20.1% of the minors failed to correctly answer more than half of the questions designed to assess basic digital skills. This percentage was higher among boys than girls, indicating a gender gap in digital proficiency that disadvantaged boys. Consistent with PISA findings, the incidence of digital educational poverty declined significantly as parental education increased.

When woven together, these multi-layered data points reveal a cumulative and reinforcing chain of evidence. The structural socio-economic and regional fractures historically documented by PISA and INVALSI provide the baseline context; the macro-level international data from ICILS exposes the specific, pre-existing deficit in digital proficiency;

⁸The “Autovalutazione di base delle Competenze Digital” (AbCD) questionnaire was administered to a sample of 772 children who attend the last class of lower secondary school, in 11 Italian cities and provinces: Ancona, Chieti, Mestre, Milan, Naples, Udine, Palermo, Rome, Turin, Velletri, Sassari.

and the micro-level findings from Save the Children illustrate how these combined vulnerabilities materialized into immediate operational failures during remote interactions. Ultimately, this evidentiary trajectory demonstrates that the pandemic did not generate new inequalities but rather acted as an accelerant—weaponizing pre-existing digital educational poverty to disproportionately penalize students who were already at the margins of the Italian education system.

3.2 Evidence from AlmaDiploma on distance learning

In spring 2020, AlmaDiploma, an association of upper secondary schools, in collaboration with the AlmaLaurea inter-university consortium, conducted a comprehensive survey of students in their final and penultimate years of upper secondary education to capture their real-time experiences with distance learning during the pandemic. The survey gathered responses from more than 70,000 students, yielding a final analytical sample of 23,305 respondents (a 32% response rate).

Methodologically, because the questionnaire was administered via Computer-Assisted Web Interviewing (CAWI), the sample exhibits an unavoidable selection bias against the most socio-economically disadvantaged students. Since completing the survey required active ICT access, the respondents inherently belonged to households that were already digitally included. Consequently, the survey’s findings must be interpreted as an optimistic upper bound of the Italian student experience. From an epistemological standpoint, this selection bias can be leveraged as a robust “pessimistic lower bound” for the systemic evaluation of educational disruption. If severe pedagogical friction, track-based stratification, and consolidated learning losses significantly emerge even within an over-represented, digitally affluent student cohort, it logically and statistically follows that the systemic disruptions suffered by highly marginalized, non-responding students were drastically more severe, validating the cumulative nature of the deprivation.

Subject to this analytical boundary, the survey evaluates three core dimensions: remote teaching tools, study load and organization, and perceived study effectiveness.

Remote teaching tools. The vast majority of respondents (93.6%) stated they received no institutional hardware or connectivity support from their schools, relying exclusively on private family resources. While 56.6% of students possessed a personal, non-shared device, this access was deeply stratified by educational track: it peaked at 59.5% among students in academic high schools (*licei*) but dropped to 43.8% among those in vocational institutes (*istituti professionali*). This indicates that personal device ownership was significantly more prevalent among older, upper-secondary students than the baseline suggested by the general ISTAT estimate of 6.1%, which covers the entire, younger student population.

Withal, a critical 4.9% of respondents reported having no access to a large-screen device (PC or tablet), forcing them to participate in remote schooling entirely via smartphone. This hardware deficit was heavily skewed by track, ranging from a marginal 3.3% in academic high schools to a staggering 15.3% in vocational schools, illustrating how institutional stratification converted family background into immediate tech-spatial limitations.

In terms of connectivity, 79.7% of students rated their connection quality as good or very good, with academic tracks maintaining a slight advantage (81.4%). Despite the widespread necessity of sharing ICT tools within the household, 89.6% of respondents reported having overall adequate equipment to attend lessons, a figure that again varied from 91.4% in academic tracks to 82.9% in vocational ones. Operationally, 72.5% of students attended daily distance learning that covered at least half of their regularly scheduled school hours. Concurrently, only 38.4% of respondents stated that their parents were able to play an active supportive role (13.1% answering “definitely yes” and 25.3% “more yes than no”). This confirms that roughly one in three older students required stronger structural support for effective learning—a proportion bound to be drastically higher among younger, less autonomous primary and lower-secondary pupils.

Study load and teaching organization. Regarding the perceived volume of work, 79.6% of students reported a substantial increase in assignments compared to traditional face-to-face instruction. Within this group, 24.7% deemed the expanded workload unsustainable, while 54.8% felt that, although heavier, it remained manageable. Conversely, the institutional response of schools was evaluated positively: 77.6% of respondents praised the organizational efficiency of their schools in planning the transition (22.5% very positively and 55.2% moderately positively). Furthermore, the pedagogical continuity provided by instructional staff achieved an overwhelming 90.6% satisfaction rate, highlighting the extensive emergency mobilization of Italian teachers.

Learning effectiveness, school tracks, and unequal educational consequences

The final segment of the survey exposes the structural limitations of the digital medium regarding actual learning consolidation. A striking 75.1% of students reported significantly lower attention spans and higher rates of distraction during remote lessons compared to physical classrooms. Only 29.6% judged remote teaching to be effective for mastering entirely new academic topics, and a mere 16.7% believed it improved their overall comprehension of the subjects relative to face-to-face instruction. Accordingly, the vast majority of students (72.1%) concluded that their final academic preparation via distance learning was inferior to what they would have achieved in person, an assessment that held true regardless of the educational track.

Looking toward the future, only 31.6% of respondents felt it would be useful to retain distance learning in a blended format after the resolution of the COVID-19 emergency. Beyond academic performance, the survey highlights a profound disruption in the non-cognitive and relational dimensions of schooling. The sudden erosion of peer-to-peer socialization and physical institutional routines led to widespread psychological fatigue, with a vast majority of students reporting feelings of isolation.

Crucially, this relational deficit was accompanied by severe anxiety regarding future employability, an issue that once again exposed the structural divide between educational tracks. While academic high-school students faced delays but maintained a clear trajectory toward university education, vocational students experienced a drastic suspension of practical laboratory activities and mandatory school-to-work internships (*Percorsi per le Competenze Trasversali e per l'Orientamento* - PCTO). For these students, who rely on immediate post-diploma labor market insertion, the lockdown did not merely disrupt their lessons—it froze the vital networks and hands-on training essential for their transition into employment, significantly dampening their professional expectations and acting as the final act of the cumulative educational divide.

In summary, while the AlmaDiploma survey paints a commendable picture of the structural resilience and adaptability of Italian schools and teachers during a crisis, it simultaneously lays bare the severe pedagogical friction of remote teaching. When these insights are paired with macro-data—such as the 8% of students who completely dropped out of remote lessons during the first wave (ISTAT, 2021)—the long-term threat to the educational capital of this generation becomes clear. Although longitudinal tracking for Italy is ongoing, international econometric simulations suggest substantial, compounding learning losses directly attributable to prolonged school closures (Hanushek & Woessmann, 2020; Azevedo et al., 2021). Crucially, the AlmaDiploma data confirms that these losses are not distributed equally: they hit vocational tracks and resource-constrained households the hardest, proving that remote teaching acted as a mechanism that converted pre-existing socio-economic inequalities into permanent educational deficits.

3.3 Dropout and educational poverty

The final component of the student-centred section links the emergency phase of remote teaching to the longer-term risk of educational deprivation. Dropout and educational poverty are read here as outcomes of a cumulative mechanism in which unequal material preconditions, unequal participation in remote schooling, and unequal learning effectiveness translated into differentiated risks of disengagement, underachievement, and school abandonment. Learning difficulties and school failure may generate educational poverty and social exclusion through different forms of dropout. In statistical terms,

two dimensions can be distinguished. The first is *explicit dropout*, commonly measured by the Early Leavers from Education and Training (ELET) indicator, which captures the share of 18–24-year-olds who have attained at most a lower secondary qualification and are no longer in education or training. The second is *implicit dropout*, which refers to students who remain enrolled or formally complete their schooling but fail to reach minimum proficiency levels in basic competences. In this sense, the share of low achievers provides an important indication of a less visible form of educational fragility.

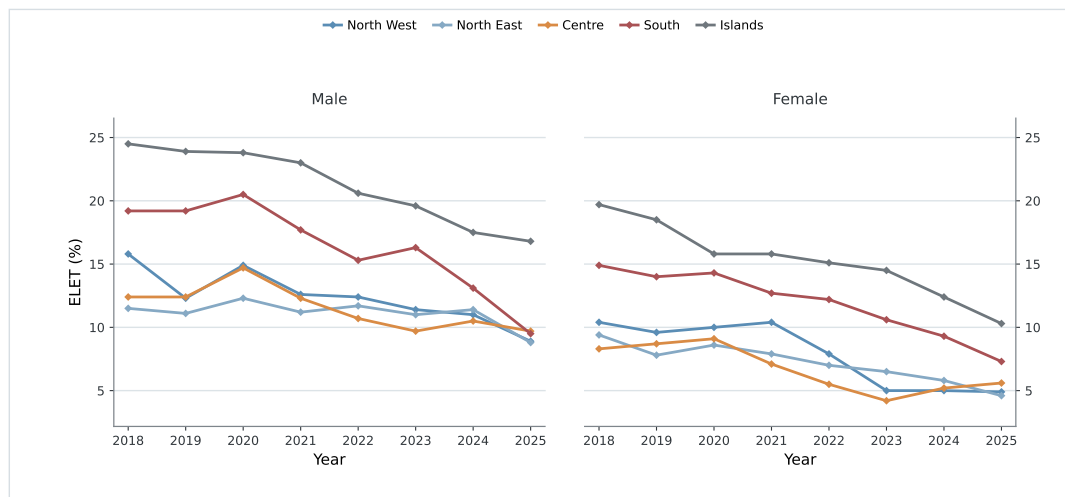
The historical territorial cleavage emerges once again in these indicators. The South and the Islands have traditionally recorded the highest rates of early school leaving, while dropout is systematically lower among females across all macro-areas. However, young women in southern regions still remain at a disadvantage compared with their peers in northern and central Italy. In light of the high share of upper secondary students who completely disengaged from remote platforms during the crisis, the risk that temporary educational detachment would turn into permanent dropout was particularly severe. Save the Children (2021) estimated that approximately 34,000 additional students were placed at immediate risk of formal school abandonment due to prolonged educational detachment during the lockdown phases.

In the following years, however, the ELET indicator followed a clear downward trajectory in Italy. As shown in Figure 3, the share of 18–24-year-olds leaving education and training early decreased from 11.5% in 2022 to 9.8% in 2024, progressively approaching the European Union average. More recent figures indicate a further decline to 8.2% in 2025, placing Italy below the European benchmark set for 2030. This aggregate improvement, however, should not be interpreted as the disappearance of educational inequalities. Figure 3 shows that ELET rates decreased after 2020 for both males and females, but with different intensities across territories and gender groups. Male early leaving remains systematically higher than female early leaving in all macro-areas, with the highest values concentrated among male students in the Islands and the South. The post-pandemic reduction in explicit dropout therefore coexists with enduring forms of vulnerability, which remain embedded in specific social and territorial contexts.

A less visible but equally relevant dimension of post-pandemic educational fragility is captured by the share of low achievers. Figure 4 reports the percentage of 15-year-old students who fail to reach baseline proficiency levels in mathematics in Italy and in the EU-27.

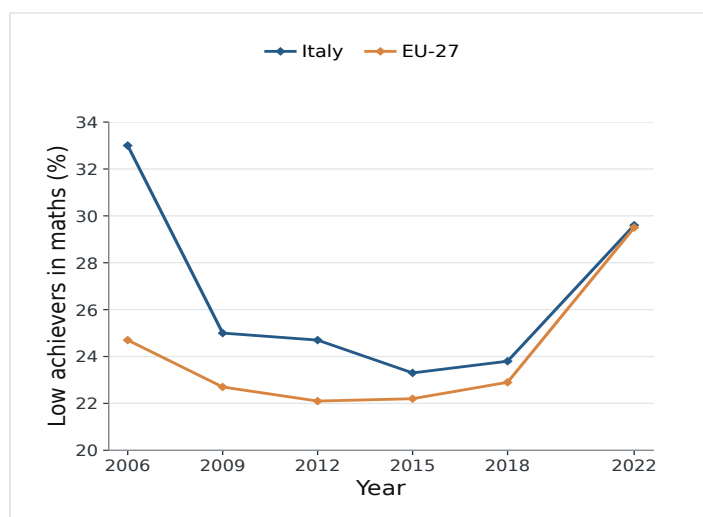
After a decline between 2006 and 2012, the share of low achievers in Italy increased from 2015 onwards, remaining consistently above the EU-27 average. The 2022 data are particularly relevant because they capture the deterioration in basic skills after the Covid-19 disruption. As Figure 4 shows, Italy experienced a marked increase in the proportion of

Figure 3: Early Leavers from Education and Training by gender and Italian macro-area



Source: Author's elaboration based on Eurostat data.

Figure 4: Low achievers in Mathematics in Italy and the EU-27, 2006–2022



Source: Author's elaboration based on Eurostat data.

students performing below the minimum proficiency level in mathematics. This worsening does not simply reflect a temporary decline in test scores; it signals a broader weakening of students' learning foundations and may increase the risk of disengagement from school. The comparison with the EU-27 average reinforces this point. Although the increase in low achievement after the pandemic is visible at the European level, the Italian trend is especially problematic because it adds to pre-existing territorial, social, and school-track inequalities.

The figure therefore helps clarify why the consequences of remote teaching cannot be assessed only through formal dropout indicators. Even where early school leaving later declined, the deterioration in basic skills points to a hidden dimension of educational deprivation that may continue to affect students' trajectories over time. A gender-disaggregated interpretation of low achievement is limited by the lack of comparable EU-level data by gender in the dataset used here. This prevents a systematic comparison between Italy and the EU-27 along this dimension. Nevertheless, the Italian data show a clear gender pattern in mathematics: male students consistently display lower shares of low achievement than female students. This suggests that the post-pandemic deterioration in basic mathematical skills intersects not only with territorial and socio-economic inequalities, but also with gender differences. However, this pattern is subject-specific. In reading, the relationship is reversed, with female students performing better than male students, while in science gender gaps are much less pronounced and in both directions. Educational vulnerability should therefore be understood as multidimensional, varying across territories, social groups, gender, and domains of competence.

Overall, this evidence indicates that the Italian school system entered the pandemic crisis with a dual structural fragility: a visible problem of early school leaving, especially in the South and Islands, and a less visible but widespread crisis of underachievement. The pandemic did not create these vulnerabilities from scratch; rather, it exposed and intensified them, contributing to a broader socio-technical polarization of educational opportunities.

4 Concluding remarks

Digital infrastructure has ceased to be a mere auxiliary asset; it now constitutes one of the structural conditions of educational participation. The Italian experience during the COVID-19 pandemic shows that digital education can create new opportunities only when the material and cognitive preconditions of access are broadly secured. When these preconditions are unevenly distributed, remote teaching does not simply compensate for the temporary loss of in-person schooling. It can instead reproduce, intensify, and

legitimize pre-existing inequalities.

The first analytical section showed that the material preconditions of remote teaching were deeply unequal before the pandemic. Italy entered the crisis with a weak position in European digital readiness, especially in human capital and digital skills, and with marked territorial gaps in broadband connectivity and device availability. These infrastructural inequalities were reinforced by domestic learning environments: overcrowded housing, the absence of a quiet study space, unequal family support, and the difficulties faced by pupils with disabilities converted the home into a highly differentiated educational setting. In this sense, digital exclusion was never only a question of internet access. It was a broader socio-material condition shaped by infrastructure, household resources, space, time, and care.

The second analytical section showed that these preconditions translated directly into unequal student experiences and outcomes. Before the pandemic, Italian students already displayed marked differences in achievement, functional literacy, and digital skills, with clear socio-economic and territorial gradients. During the first lockdown, participation in remote teaching depended heavily on private household resources, personal devices, school-track location, and students' capacity to sustain attention and autonomy in a digital environment. The evidence from upper secondary students indicates that even relatively digitally included respondents experienced increased workloads, lower attention, weak perceived effectiveness for learning new content, and strong differences between academic and vocational tracks. These mechanisms help explain why the crisis was not limited to a temporary interruption of instruction, but became part of a wider process of educational poverty, explicit dropout, and implicit dropout.

The cumulative mechanism identified in the paper can be summarized through the relationship between the two analytical sections. The section on material preconditions shows why access to remote teaching was unequal from the outset: connectivity, devices, housing space, and family support were distributed according to territorial and socio-economic lines. The section on students shows how these unequal starting points became unequal educational consequences: weaker participation, reduced learning effectiveness, stronger difficulties in vocational tracks, and higher risks of disengagement and educational poverty. The transition from the first section to the second is therefore also the transition from inequality of access to inequality of outcome.

In conclusion, the Italian case demonstrates that digital innovation in education cannot be evaluated only in terms of technological availability or institutional adoption. Its social effects depend on the distribution of the resources that allow students to use technologies meaningfully. When a school system relies on remote platforms without ensuring universal access to devices, reliable connectivity, adequate study spaces, family or institutional

support, and functional digital literacy, digital education risks becoming an institutional mechanism that magnifies learning losses and reduces social mobility. Addressing this risk requires a coordinated policy shift away from the mere procurement of technology and toward the structural reduction of educational poverty. Investments in high-quality ICT resources, digital pedagogical training, targeted support for vulnerable territories, inclusive provision for students with disabilities, and specific measures for vocational tracks should therefore be understood not as emergency responses, but as democratic preconditions for an equitable digital future of education.

References

- AlmaDiploma (2020). *Indagine sulla Didattica a Distanza (Rapporto)*.
https://www.almadiploma.it/info/pdf/indagini/altro/2020/AD_IndagineSullaDidatticaADistanza.pdf
- Andrew, A., Cattan, S., Costa Dias, M., Farquharson, C., Kraftman, L., Krutikova, S., Phimister, A., & Sevilla, A. (2020). Inequalities in Children's Experiences of Home Learning during the COVID-19 Lockdown in England. *Fiscal studies*, 41(3), 653–683.
doi:[10.1111/1475-5890.12240](https://doi.org/10.1111/1475-5890.12240)
- Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A., and Geven, K. (2021). Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates. *The World Bank Research Observer*, 36(1), pp. 1-40.
- Bovini G., and De Philippis M. (2021). *Alcune evidenze sulla modalità di svolgimento della didattica a distanza e sugli effetti per le famiglie italiane*. Note-Covid 19, Banca d'Italia.
- Engzell, P., Frey, A., and Verhagen, M. D. (2021). Learning Loss Due to School Closures during the COVID-19 Pandemic. *Proceedings of the National Academy of Sciences*, 118(17), e2022376118.
- Fasanella, A., Dentale, M., Capozza, V., and Rimoli, G. V. (2025). Il contributo della didattica digitale allo sviluppo dell'apprendimento per competenze: luci e ombre. In V. Lo Presti and M. Dentale (eds.), *La valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders*, FrancoAngeli, Milan, pp. 80-128.
- Fasanella, A., and Faggiano, M. P. (2022). Italian Secondary School Students and the Distance Learning Experience: From Current Critical Issues to Future Opportunities, Reflecting on VR. *Italian Journal of Sociology of Education*, 14(3), pp. 17-44.
- Fraillon, F.J., Ainley, J., Schulz, W., Friedman, T., and Duckworth, D. (2020). *Preparing for Life in a Digital World. IEA International Computer and Information Literacy Study 2018 International Report*, Springer International Publishing.
- Hanushek, E.A., and Woessmann, L. (2020). *The Economic Impacts of Learning Losses*, OCSE library:
<http://www.oecd.org/education/The-economic-impacts-of-coronavirus-covid-19-learning-losses.pdf>

Hodges, C., Moore, S., Lockee, B., Trust, T., and Bond, A. (2020), *The Difference between Emergency Remote Teaching and Online Learning*. EDUCAUSE Review, 27 March.

<https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>

INVALSI (2020). *ICILS 2018: Rapporto nazionale. Competenze digitali degli studenti italiani nel contesto internazionale*. Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione.

ISTAT (2021), *Rapporto annuale 2020*,

<https://www.istat.it/storage/rapporto-annuale/2020/Rapportoannuale2020.pdf>

Lo Presti, V., and Dentale, M. (2021). La valutazione dell'impatto sociale della didattica a distanza nelle scuole secondarie di secondo grado romane: istanze, esigenze e percorsi di cambiamento. *RIV. Rassegna Italiana di Valutazione*, 80-81, pp. 124-146.

Lo Presti, V., and Dentale, M. (eds.) (2025), *La valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders*, FrancoAngeli, Milan.

Lo Presti, V., Capozza, V., Salvi, V., and Cavallo, S. (2025). Stili di didattica e Pratiche pedagogiche a confronto: contesti e rappresentazioni. In V. Lo Presti and M. Dentale (eds.), *La valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders*, FrancoAngeli, Milan, pp. 55-79.

Lo Presti, V., Faggiano, M. P., and Dentale, M. (2025), Relazioni sociali, istanze e bisogni socio-emotivi durante e dopo la pandemia. In V. Lo Presti and M. Dentale (eds.), *La valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders*, FrancoAngeli, Milan, pp. 129-170.

Martire, F., Palmieri, M., Mastrolia, A., and Seller, D. (2025). Caratteristiche e specificità della didattica digitale in tempi di pandemia: uno zoom sui contesti scolastici coinvolti nella ricerca. In V. Lo Presti and M. Dentale (eds.), *La valutazione d'impatto sociale della didattica digitale dopo il Covid-19. Una prospettiva di ricerca multi-obiettivo e multi-stakeholders*, FrancoAngeli, Milan, pp. 29-54.

OECD (2020), *Learning remotely when schools close: How well are students and schools prepared? Insights from PISA*, OECD Publishing.

<https://www.oecd.org/coronavirus/policy-responses/learning-remotely-when>

-schools-close-how-well-are-students-and-schools-prepared-insights-from-pisa-3f9dd1e7/

OECD (2021), *21st-Century Readers: Developing Literacy Skills in a Digital World*. PISA OECD Publishing.

https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/05/21st-century-readers_26f2b462/a83d84cb-en.pdf

Potter, J., and McDougall, J. (2017), *Digital Media, Culture and Education: Theorising Third Space Literacies*, Palgrave Macmillan.

Sarno, E. (2022). Innovation Technology and Education: New Opportunities or New Inequalities? The Italian Experience. *Journal of Performance Management*, 1(1), pp. 41–54. ISSN: 2820-7327.

Save the Children (2020), *Riscriviamo il futuro: L'impatto del coronavirus sulla povertà educativa*, disponibile su:

https://s3.savethechildren.it/public/files/uploads/pubblicazioni/limpatto-del-coronavirus-sulla-poverta-educativa_0.pdf

Save the Children (2020), *Riscriviamo il futuro: Rapporto sui primi sei mesi di attività*, disponibile su:

https://s3.savethechildren.it/public/files/uploads/pubblicazioni/riscriviamo-il-futuro-rapporto-6-mesi_1.pdf

Save the Children (2021), *Riscriviamo il futuro: Una rilevazione sulla povertà educativa digitale*, IPSOS per Save the Children, disponibile su:

https://s3.savethechildren.it/public/files/uploads/pubblicazioni/riscriviamo-il-futuro-una-rilevazione-sulla-poverta-educativa-digitale_0.pdf

Discrimination of immigrants in mortgage pricing and approval: Evidence from Italy*

Paolo Emilio Mistrulli[†] Md Taslim Uddin[‡] Alberto Zazzaro[§]

Abstract

In this paper, we explore empirically whether immigrants, other things being equal, pay more for mortgages than natives and whether the probability that banks approve their loan applications is systematically lower. To this aim, we use two extensive and unique dataset of mortgage contracts and banks' requests for initial information about potential mortgagors drawn from the Italian Credit Register for the period 2011-2016, and survey data from the Survey on Household Income and Wealth conducted by the Bank of Italy for the period 2006-2016. We find that immigrants pay 20-24 basis points more than native Italians on single-name mortgages and 28-40 basis points more on jointly-owned ones. This interest rate gap narrows significantly, but does not disappear, when immigrant borrowers' credit history lengthens or if they borrow from a cooperative bank. Finally we find that immigrants have a 2.7% smaller chance of getting a mortgage compared to natives, which decreases for mortgage applications submitted to cooperative banks. Overall, our findings suggest that the disparity of treatment of immigrants in the Italian mortgage market is mostly due to a greater difficulty of banks in assessing the credit-worthiness of culturally distant borrowers. However, we also detect that cultural distance may fuel persistent disparity between migrants and natives.

Keywords: Immigrants; discrimination; mortgage lending; interest rates; loan approval.

JEL Codes: G21, J15, J71

*We are grateful to Giulia Bettin, Luca Coraggio and Tommaso Oliviero for helpful discussion and comments. The opinions expressed in this study are our own and do not necessarily represent those of the Bank of Italy.

[†]Bank of Italy. E-mail: paoloemilio.mistrulli@bancaditalia.it

[‡]Corresponding author. Independent University Bangladesh. E-mail: mtaslim@iub.edu.bd

[§]University of Naples Federico II, CSEF and MoFiR. E-mail: alberto.zazzaro@unina.it

1 Introduction

Discrimination in mortgage lending against immigrants and ethnic minorities has been an important and longstanding issue in academic and policy debate (Ladd, 1998; LaCour-Little, 1999; Turner and Skidmore, 1999). Access to mortgages at affordable and fair interest rates is a major barrier to homeownership (Henderson and Ioannides, 1983; Haurin et al., 1997; Guiso and Jappelli, 2002; Chiuri and Jappelli, 2003; Bayer et al., 2014; Trucchi, 2016; Goodman et al., 2018) and a crucial step towards actual and perceived socio-economic integration of immigrants in the host countries (Borjas, 2002; Constant et al., 2009; Arbaci and Malheiros, 2010; Accetturo et al., 2014; Kara and Molyneux, 2017; Bertocchi et al., 2022).

In this paper we empirically analyze immigrants' access to credit in the Italian mortgage market. We use Bank of Italy's Credit Register (CR) data on mortgages originated by commercial banks operating in Italy in the period 2011-2016 and on banks' requests of initial information about borrowers to the CR in 2015 and 2016. Our main focus is on mortgage pricing: we explore whether immigrants pay more for mortgages than Italian natives and whether the possible different treatment reflects the different credit qualification of mortgagors or can also be due to forms of discrimination by banks. In addition, we complement the interest rate analysis with evidence on the probability of mortgage approval by Italian banks for loan requests from immigrants and natives.

Economic theory provides two well known explanations for discrimination against minorities, based on taste or lack of information between the trading parties (Becker, 1971; Phelps, 1972; Arrow, 1973). In the first case, discrimination in mortgage markets takes place because of personal prejudice or preferences of lenders against a certain minority group. In the second case, unbiased native lenders can use race or ethnicity as a proxy for individual unobserved factors, like family links and social network, that have an impact on credit risk. In this way, due to the lack of complete information about individual borrowers' risk, lenders end up treating differently two otherwise identical native and immigrant mortgagors. At the turn of the two explanations based on preference and information, an additional cause for discrimination can be the cultural distance between native lenders and

immigrant mortgagors. Lack of cultural kinship makes it difficult for lenders to understand informational cues about immigrant creditworthiness (Cornell and Welch, 1996) and, at the same time, makes native-born loan officers less likely to empathize with borrowers of an ethnic other than their own and to help them borrow on fair price and non-price terms (Kim and Squires, 1998; Bostic, 2003; Fisman et al., 2017; Albareto et al., 2022; Accetturo et al., 2023).

However, testing for discrimination and disentangling the nature of discrimination in mortgage lending by using real data is especially challenging.¹ First, mortgages are long term and multiple lending from different bank are rare. This implies that a single interest rate figure per borrower is available and it is not possible to control for borrower unobservable characteristics. Second, bank lending policies can provide for formal or informal rules, like minimum mortgage amount cutoffs, requirement for a personal guarantee by a third party or other redlines, which systematically make immigrant borrowers at a disadvantage based on some common characteristics unobservable to econometricians, like family wealth or social kinship. In this case, higher interest rates on mortgages to immigrants or lower probability of loan approval might hide systematic differences between immigrant and native-born borrowers. Third, banks offer borrowers a menu of mortgage terms to choose from, including interest rates, loan down payments, upfront fees and others, which can be not available to researchers. Once again, in a similar context, higher interest rates on mortgages to immigrants might capture their unobservables preferences or characteristics rather than a discrimination in the menu of conditions offered by the banks to immigrants and natives.

In our sample, the unconditional mean interest rate on mortgages made out to immigrants is statistically significantly higher than that on mortgages to Italian natives by 27 basis points (bps); when we distinguish between single-name and joint mortgages this difference is 18 and 32 bps, respectively. To investigate whether this gap reflects discrimination against immigrants and what its possible nature is, we perform various econometric

¹Some studies have used an experimental setting by using paired testing and email-correspondence methods to compare the response of loan officers to mortgage requests by otherwise identical white and minority applicants, finding evidence of discrimination against the latter group Ross et al. (2008); Hanson et al. (2016).

exercises. First, we condition the effect of immigrant status on mortgage interest rates on observable mortgagors' characteristics and risk factors included in our dataset, and on a set of bank, province and time fixed effects. The disparity in interest rates paid by immigrants and natives remains statistically significant, ranging from 24 to 40 bps for single-name and joint mortgage contracts. Unfortunately, credit register does not contain information about the loan down payments and the income of the mortgagor, two key risk factors that are commonly used by banks to fix the interest rate on mortgages that can systematically differ between immigrants and natives (Elmer and Seelig, 1999; Elul et al., 2010). To address this concern, we use a machine learning approach to generate imputed values for mortgage loan-to-value (LTV) ratios and individual income from survey data reported in the "Survey on Household Income and Wealth" (SHIW) conducted by the Bank of Italy, and we again find that immigrants pay significantly more than native-born Italians (20 bps on single-name mortgages and 29 bps on joint mortgages).

Then, we use the Oaxaca-Blinder decomposition method to separate the average interest rate differential paid by immigrant mortgagors into the part explained by the differences in the explanatory variables between the two groups and the part not explained by the differences in observable characteristics. We find that only 15% of the price gap can be attributed to differences in the individual risk factors and mortgage characteristics, including the imputed income and LTV, while 85% is driven by unobserved factors, consistent with possible unobserved discrimination factors against immigrants.

In addition, we use propensity score matching methods to mitigate possible selection biases and compare mortgages taken up by immigrants to native-born Italians with similar observable characteristics. Even in this case, immigrant mortgagors result to pay higher interest rates than native-born Italians.

Finally, we repeat our regression analysis using survey information on mortgages self-reported by a representative sample of households participating in the Bank of Italy's SHIW, confirming once again the existence of a statistically significant price gap on mortgages taken up by immigrants.

To shed some light on the source of price discrimination against immigrants, first, we

consider the country of origin of immigrants. In line with results of [Albareto et al. \(2022\)](#) for loans to sole proprietorship firms, we find that the interest rate gap is the highest for mortgages taken by Asians, Africans and South-Americans, which can be considered the immigrant groups most culturally distant from native-born Italian loan officers, and it is the lowest for mortgages to North American mortgagors. Second, in order to distinguish between statistical- and taste-based discrimination, we look at three sources of heterogeneity. We consider the mortgagors' credit history measured by the number of years since the first appearance in the central credit register. The idea is that if the interest rate gap shrinks for immigrants with a longer credit history, this indicates that the price gap is likely driven by the fact that banks have less information on the unconditional probability of repayments for the group of immigrant mortgagors. Then, we distinguish between dominant bank in the market of mortgages and fringe competitors, and between cooperative and non-cooperative banks. In the first case, the assumption is that large competitors are more likely to take their lending decisions based on automated credit scoring models and statistical information on borrower groups. On the contrary, cooperative banks are expected to have a more in-depth knowledge of the local economy and society and rely on borrower-specific soft information. Our regression results partly support the relevance of statistical-based discrimination, showing that as length of credit history increases the interest rate gap for immigrant mortgagors decreases statistically significantly, as well as for immigrants borrowing from cooperative banks, while we do not find any difference between mortgages granted by dominant and fringe competitors. In any case, although smaller, the effect of immigrant status on the interest rate remains positive and significant regardless of the different availability of specific information that banks have on average on the borrower.

Lastly, we complement the analysis of discrimination in interest rates with evidence on mortgage approvals. Although we do not observe individual loan applications and their final outcome at the approached banks, the Italian CR contains data on the requests for initial information about potential borrowers submitted by each bank to the CR and the loans that they actually approved in the following months. Therefore, using the approach proposed by [Jiménez et al. \(2012\)](#), we analyze the probability that the request for initial

information about a borrower made by a bank will be followed by the disbursement of a loan to that borrower by the same bank within a certain span of time. Also regarding loan approvals, regression results confirm that immigrants have more restricted access to credit, showing that they are statistically significantly less likely to receive the mortgage (precisely, immigrants have a 2.7% smaller chance of getting a mortgage compared to natives). Once again, the discrimination effects is significantly lower for loan applications to cooperative banks, while credit history does not seem to have positive effects.

Unequal treatment of mortgage applicants based on race, ethnic and national origin has long been empirically documented for the United States, examining all the lending stages, from mortgage processing (Wei and Zhao, 2022) to mortgage performance (Reid et al., 2017). Studies on mortgage pricing by US banks consistently find that, after controlling for bank, mortgagors, loan and property characteristics, ethnic minorities (African-, Asian- and Hispanic-Americans) are both more likely to borrow at unfavorable terms and pay on average higher interest rates than White Americans. This holds independent of considering average annual rates or overages, prime or subprime mortgages, male or female mortgagors (Black et al., 1978; Courchane and Nickerson, 1997; Black et al., 2003; Boehm et al., 2006; Boehm and Schlottmann, 2007; Bocian et al., 2008; Bayer et al., 2014; Ghent et al., 2014; Cheng et al., 2015; Bayer et al., 2017; Delis and Papadopoulos, 2019; Ambrose et al., 2021; Bartlett et al., 2022).²

Unlike the United States, where anti-discrimination laws - such as the Fair Housing Act of 1968, the Equal Credit Opportunity Act of 1974 and the Home Mortgage Disclosure Act of 1975 - have long been enacted to protect ethnic minorities from discrimination in the real estate and in the mortgage market, neither Italy nor the rest of Europe has specific laws to ensure individuals and businesses fair access to the loan markets. The lack of law dealing with this issue is matched by the almost complete lack of empirical studies on the disparity of mortgage pricing against immigrants in European countries. A notable exception is Diaz-

²Two exceptions are Crawford and Rosenblatt (1999) and Bhutta and Hizmo (2021). The former explores mortgage prices charged by a major mortgage lender in the US finding that they are broadly race-neutral. Bhutta and Hizmo (2021), document that although Black- and Hispanic-American borrowers pay higher interest rates, they also pay lower up-front costs, consistent with the absence of discrimination against minorities in the multidimensional mortgage pricing and the presence of differences in preferences that lead minorities to choose a different location in the fee-rate schedule than whites.

Serrano and Raya (2014) who analyze a large sample of mortgages to homebuyers mediated by the mortgage brokerage branch of a real estate company in Spain. They find that immigrants pay 18 basis points higher interest rates for mortgages than do native even after controlling for differences in credit worthiness and other factors.³ With regard to Italy, Albareto et al. (2022) is the only study of which we are aware that analyzes possible discrimination of immigrants in accessing credit. They look at pricing of lines of credit to sole proprietorship firms. By using central credit register data for the period 2004-2008, they find that interest rates paid by immigrant-owned enterprises are on average 36 basis points higher than those paid by enterprises owned by native-born Italian entrepreneurs, and that this difference is greater for immigrants from Asian and non-Christian countries.⁴ We add to their analysis by considering, for the first time in the literature, the pricing of home mortgages to immigrant households.

The rest of the paper is organized as follows. Section 2 presents the context of immigration and mortgage market in Italy. Section 3 presents the data sources, the dependent and independent variables and the empirical strategy. Section 4 discusses our regression results on mortgage contract data drawn from the Credit Register, section 5 presents results using survey data drawn from the SHIW dataset, while section 6 presents results on loan approval. Section 6 concludes.

2 The Italian context

2.1 Immigrants in Italy

Italy is a country of recent immigration. Only from the mid-seventies of the twentieth century, Italy experienced a positive migration balances. In Figure 1 we display the annual inflows of foreign population in Italy starting from 2000 as reported in the OECD International Migration Database.

³Kara and Molyneux (2017) study differential access to mortgages in the UK but do not consider interest rate: they find that low-income Black households are significantly less likely to have a mortgage compared to low-income White households.

⁴Few other studies have examined gender discrimination in access to credit in Italy without considering the immigrant status of the borrower (Bellucci et al., 2010; Alesina et al., 2013; Stefani and Vacca, 2013; Calcagnini et al., 2015).

[FIGURE 1 AROUND HERE]

In 2007 there was a peak in the entry of new immigrants, with over 510,000 people. After that date, also due to the financial crisis and the strong slowdown of the Italian economy, the annual inflow of immigrants decreased, settling at around 300,000 people per year. In 2018, more than 6 million foreign-born people reside in Italy legally, constituting 10.4% of Italy's total population.

During the years considered in our analysis, Romanians and Albanians are the two most important immigrant communities in Italy, followed by the Moroccan, Chinese and Ukrainian ones (see Table 1). However, there have been substantial changes in the composition of immigrants by nationality with a sharp increase in the number of immigrants from Asian and African countries – Bangladesh (+44%), Nigeria (+44%) and Pakistan (+34%) – and a reduction of immigrants from East Europe – Macedonia (-22%) and Poland (-10%).

[TABLE 1 AROUND HERE]

The housing conditions of foreign-born population in Italy are on average rather poor, with almost 40% living in overcrowded or substandard dwellings (compared to an average share of almost 20% in other European Union countries), while the share of the foreigners who perceive themselves as discriminated against on the grounds of ethnicity, race, or citizenship is 14%, not different from the European average.⁵

2.2 Italian mortgage market

The mortgage market is a major driver of developed economies. In 2019, the total outstanding residential loans in the eurozone plus the United Kingdom is 54% of the area's GDP, while the amount of new residential loans (consisting of new mortgage loans plus remortgaging with another bank) is 1.19 trillion of euros, equal to 8.2% of GDP (EMF-ECBC, 2021).

Compared to the average for European countries, the mortgage market in Italy is significantly smaller, with a stock of outstanding residential loans and a flow of new mortgages

⁵Data are from OECD *Indicators of immigrant integration*, available at: <https://www.oecd.org/els/mig/indicatorsofimmigrantintegration.htm>

that are equal to, respectively, 21% and 6% of GDP. Although, Italy is characterized by a high home-ownership rate, households make less use of the mortgage for the purchase of the house. According to the OECD statistics, 71% of Italian dwellings were occupied by owners in 2018, well above the homeownership rate registered in France (60,6%), Germany (43,8%) and the United Kingdom (64,9%). However, made 100 the number of houses occupied by owners, only 14% were occupied by household with a mortgage in place, compared to 61% in the United States, 47% in the United Kingdom, 42% in Germany and a European average of 28%. In 2020, only half of the more than five hundred thousand housing transactions negotiated in Italy are with a mortgage. These figures reflect the lower propensity to borrow of Italian households, whose total debt is 87% of their disposable income in 2018, a value significantly lower than that of France (120%), Germany (93%), Spain (107%), the UK (145%) and the United States (103%) (Guiso et al., 2022).

[FIGURE 2 AROUND HERE]

The average mortgage amount provided by Italian banks in 2020 is 130,000 euros, compared to €277,000 in Germany, €240,000 in the UK and €230,000 in the US (EMF-ECBC, 2021). The cautious stance of Italian banks and households is reflected in the loan-to-value ratio (LTV) of mortgage loans, which is 63% in Italy, and in the share of new loans with an LTV greater than 80% which does not exceed 16%, both values much smaller than the European average of 81% and 56,1%, respectively (Lang et al., 2020).

The two types of mortgage contract offered by Italian banks are the adjustable-rate mortgage (ARM), where the interest rate charged to the borrowers provides for a spread on an underlying reference rate (usually, 1-3 months Euribor), and the fixed-rate mortgage (FRM), where interest rate and installments are fixed for the whole length of the mortgage. In our sample period, just over 28% of the mortgages issued are FRMs. In figure 3 we report the average interest rate charged on FRMs and ARMs. The former has been higher throughout the period 2011-2016, a period of decreasing interest rates (plot (a)). This is true both for mortgages to Italian natives and for those taken up by immigrants, even if the latter pay higher rates independent of the type of contract (plot (b)).

[FIGURE 3 AROUND HERE]

3 Data, variables and statistics

3.1 Datasets

Our main analysis on the pricing of mortgages is at the loan level. We draw data from the Central Credit Register (CR) and the Analytical Survey on Loan Interest Rates (ASLIR) maintained by the Bank of Italy.⁶ From the ASLIR we obtain the contractual interest rates on personal loans granted to natural persons and sole proprietorships by a subset of banks in each quarter, including all the main banking groups active in Italy which cover more than 90% of the Italian credit market. Precisely, the reporting banks are required to submit information for each new loan made in a reference quarter to borrowers for whom, at the end of that same quarter, the sum of all the accorded and used loans reported in Central Credit Register is equal to or more than 75,000 euros. The ASLIR also includes data on the identity of the bank originating the loan, the type of contract, the loan size at origination and a number of personal data of the borrower. We complement this information with the date on which the borrower first appears in the Bank of Italy's credit information system obtained from the CR to measure her/his credit history length.

Our dataset includes all the home mortgages in the ASLIR, both for the purchase and renovation of a dwelling, originated between 2011 and 2016. The original dataset has complete records on around 1 million mortgages. To remove outliers, data have been trimmed to the top and bottom one percentile of the interest rate distribution (which represent 19,952 observations). We also exclude the contracts for which the initial duration is less than 5 years (1,992 observations) and the loan size is less than 5,000 euros (1,033 observations).⁷ In this way we exclude from the analysis the mortgages that are more likely devoted to small home renovations. Finally, we exclude joint mortgage contracts with more than two co-borrowers to facilitate the identification of the role of the immigrant status of (co-)mortgagors (24,174 observations).⁸

⁶In Italian, the *Rilevazione Analitica dei Tassi di interesse*, also known by the acronym TAXIA (Bank of Italy, 2021).

⁷All the mortgage contracts that we consider in the analysis have a maturity of more than 5 years but data on the exact duration of each contract are not available.

⁸In the whole sample, the maximum number of co-mortgagors in a joint mortgage contract is six.

After these filtering steps, we end up with a sample of 954,341 loans relative to 24 quarters from March 2011 to October 2016, of which 490,189 single-name mortgages and 464,152 joint mortgages.

For robustness, in Section 5 we repeat our regression analysis using survey data at the household level drawn from the Survey on Household Income and Wealth (SHIW) conducted by the Bank of Italy for the period 2006-2016.

Finally, in order to build a measure of loan approval at the bank-applicant level, we use Credit Register data on requests for "CR first information" submitted by Italian banks about potential borrowers approaching their institute in the period 2015-2016. In sections 5 and 6 we provide a detailed description of, respectively, the SHIW and "CR first information" datasets and of variables used in the our empirical exercises.

3.2 Credit Register variables for mortgage pricing analysis

Table 2 contains a list of all dependent and independent variables drawn from the CR used in the analysis of mortgage pricing, their definitions, mean values and standard deviations.

[TABLE 2 AROUND HERE]

Our dependent variable is the contractual annual percentage rate charged by banks on mortgage loans (*APR*), including interest payments and any additional fees and costs associated with the transaction. On average, during our sample period the mortgage interest was 3.22%, varying from a minimum of 0.17% to a maximum of 7.73%.

The key explanatory variable of *APR* is *Immigrant*, a dummy that takes the value of 1 for non-Italian born mortgagors and 0 otherwise. In our sample 6% of mortgagors are immigrants. Of these, 70% are from an European country (the majority from Eastern Europe), 11% are Asians, 9% are South Americans, 7% are Africans and just over 1% come from North America and Oceania. Our sample includes both mortgages in the name of a single borrower and mortgages made out to two co-borrowers, indicated by the dummy *Coint* that takes the value 1 for the joint mortgages. The share of single-name mortgages taken out by immigrants – $\sum_i Singleimmi / \sum_i (1 - Coint)$ – is 5%, while the share of joint

mortgages in two immigrant names or made out to one immigrant co-borrower – given by $\sum_i Cointimmi / \sum_i Coint$ and $\sum_i Cointmix / \sum_i Coint$ – are 6.7% and 6.9%, respectively.

Furthermore, the Central Credit Register includes information on the mortgage loan amount in euros (*Loan amount*), on whether the mortgage is fixed rate or adjustable rate (*Fixed*), on the age and gender of the primary borrower, (*Age* and *Male*), and her/his *Credit history*). In our sample, the average size of a mortgage loan is €139,309 and 28% of them are fixed rate. The average age of mortgagors is 41 years and in 56% of contracts the primary mortgagor is male. The average *Credit history*, which captures the public information about borrowers available to banks, is 2.23 years long.

4 Results on mortgage pricing with CR data

4.1 Univariate analysis

In table 3 we report t-test results for mean differences in loan and borrower characteristics between mortgages made out to native-born Italians and immigrants. We consider mean differences for single-name and joint mortgage contracts (Panel A and Panel B, respectively) and for the full sample, including both types of mortgages (Panel C).

[TABLE 3 AROUND HERE]

Overall, there are statistically significant differences in all the mortgage dimensions between immigrant and native mortgagors. The unconditional mean mortgage rate paid by immigrants is 26 bps higher than that paid by natives. This gap is greater for joint mortgages (32 bps) than for mortgages made out to single borrowers (18 bps). The average amount of a mortgage jointly owned by immigrants is approximately 16,000 euros less than that granted to two Italian co-borrowers, while for single-name mortgages this difference is approximately 5,900 euros. Finally, immigrants tend to take out ARM loans more frequently than natives.

The individual characteristics of immigrant and Italian native mortgagors are also slightly different. Immigrant mortgagors are younger and female in a larger proportion compared to

native ones. Importantly, they have a shorter credit history: the share of immigrant borrowers with no credit history is 81%, compared to 65% for Italian natives, while their average credit history is almost half shorter than for natives (1.14 versus 2.30 years).

4.2 Baseline regression model

We estimate the following multivariate linear regression model of mortgage interest rate:

$$APR_{ijt} = \alpha \cdot Immigrant_j + \sum_{n=1}^N \beta_n \cdot X_{ijt}^n + \mu_j + \delta_i + \epsilon_{ijt} \quad (1)$$

where the dependent variable APR_{ijt} is the annual percentage interest rate charged on the mortgage loan granted to borrower i by bank j in quarter t , $Immigrant$ is an indicator for the immigrant status of borrower j , X_{ijt} is a set of N mortgage and borrower characteristics, δ_i and μ_j denote, respectively, a series of bank and province fixed effects, where province is the place of residence of the borrower, and ϵ_{ijt} is the error term. The estimated coefficient α represents the interest rate gap between mortgage loans to immigrant and native borrowers.

In our baseline analysis, we use an OLS estimator with robust standard errors and include in X_{ijt} only variables recorded in the ASLIR and CR. Table 4 reports the estimation results. We estimate model (1) separately for the subsample of single-name mortgages in columns [1]-[3], and for joint mortgages in columns [4] and [5]. In the last two columns [6] and [7], we show results for the full sample that includes both type of mortgages. For all the samples, first we consider a specification with only the immigrant status of the mortgagors and the bank, province and time fixed effects, then we present specification with control variables. In the Appendix, we show that our results are qualitatively and quantitatively robust to the inclusion of bank $\times$ time and province $\times$ time fixed effects, and to the use of robust standard errors clustered at the bank level (see Table A1).

[TABLE 4 AROUND HERE]

Starting from single-name mortgages, the coefficient on *Immigrant* indicates that that non-Italian borrowers pay 24 bps more on mortgages compared to Italian natives. This difference remains broadly stable once we control for observable mortgage and mortgagor

characteristics. When we take into account the immigrants' origin and substitute the dummy *Immigrant* with a set of dummies indicating their continent of origin, as in column [3], we find a statistically significant interest rate gap for all the immigrant groups. Interestingly, however, the gap tends to be greater for Asian, Central American and African immigrants (32, 26 and 25 bps, respectively), while Oceanians and North Americans are the least penalized in interest rate conditions, paying, respectively, only 10 and 6 bps more than Italian natives. These results are consistent with findings of (Albareto et al., 2022) for interest rate on credit lines to sole proprietorships owned by immigrant and Italian native entrepreneurs, and, to the extent that African, Asian and South American are the ethnic groups culturally more distant from native-born loan officers, suggest that the interest gap in mortgage pricing may be in part driven by lender-borrower cultural diversity.

The interest gap increases for joint mortgages made out to two immigrant co-borrowers that, on average, pay 40 bps more than mortgages in name of two Italian natives (column [5]). However, for immigrants borrowing together with a native borrower (*Cointmix* = 1), the gap, while remaining statistically significant, is reduced to 12 bps, indicating that the presence of a native-born co-borrower acts as a sort of guarantee for the bank and mitigates the cultural distance with immigrant borrowers. These findings are confirmed in the full sample specification [7]. Compared to single-name mortgages made out to Italian natives, the interest rates on single-name and joint mortgages to immigrants are, respectively, 25 and 49 bps higher, while the interest rate on "mixed" mortgages with immigrant and native co-borrowers is higher by 21 bps.

Moving on to the control variables, the coefficient on loan amount is negative, consistent with previous evidence suggesting that big borrowers have greater bargaining power and are able to achieve better contractual terms Black et al. (2003); Alesina et al. (2013); Albareto et al. (2022). Contractual interest rate on FRMs is about 75 bps higher than that on ARMs. Age has a U-shaped marginal effect on the interest rate, reaching its minimum value for 45 years old mortgagors, while we do not find any evidence of economically significant gender gap in interest payments on mortgages, with male borrowers paying, on average, about 2 bps more than the female ones. Finally, the coefficient on *Credit history* is negative and

statistically significant, in line with the hypothesis that applicants who are already registered in the CR at the time of taking the mortgage contract are informationally less opaque for the bank and pay lower interest rates.⁹

4.3 Imputing LTV and mortgagors' income: A machine learning approach

The CR and ASLIR do not contain information on the LTV ratio and the income of mortgagors. However, the portion of the lender-assessed value of the house that is covered by the mortgagors' own financial resources and the earning capacity of the latter are two key risk indicators used by banks to decide on the loan application and the related interest rate, which often follows an explicit LTV-interest-rate supply schedule. In addition, both risk factors – LTV ratio and earning capacity – tend to be more severe for immigrant mortgagors than for the Italian native ones.¹⁰ This introduces a possible omitted variable bias in our baseline estimations in table 4, which results in an overestimation of the interest gap towards immigrants.

To address this concern, we adopt a machine learning-based random forest approach (hereafter RF) for missing data imputation. Precisely, we use RF to predict and impute missing income and LTV values for mortgages included in the CR dataset by exploiting survey information on LTV and income contained in the Bank of Italy's SHIW dataset.

RF is a supervised machine learning algorithm proposed by Ho (1995) and Breiman (1996; 2001) for classification and regression problems that recursively identifies split points that subdivide the data into subgroups based on the values of a number of relevant feature variables (or predictors). It has been extended to address data imputation problems by developing various RF missing data algorithms (Tang and Ishwaran, 2017). In RF a large number n of independent decision trees are randomly generated from a training dataset with replacement through bootstrap sampling of k records containing a set of features X_T

⁹Results of table 4 remains qualitatively and quantitatively unchanged when we limit the analysis to mortgage loans above 75,000 euros, the threshold of borrowers' total outstanding loans considered by the ASLIR (results are reported in the online Appendix, table A2).

¹⁰This find clear confirmation in the Bank of Italy's SHIW dataset, where the unconditional mean values of LTV and income for immigrants are statistically significantly lower than for natives. Precisely, if we pool SHIW waves from 2006 to 2016, conditional on holding a mortgage, immigrants display an average LTV ratio equal to 87% and income equal to 16,018 euros compared to 61% and 28,216 euros of Italian natives (the differences between the means are both statistically significant at the standard 5% significance level).

and an output variables y_T . At each candidate split point in the training process, RF uses a random subset of the features. Finally, each tree generates an output and the final output is obtained by averaging the tree outputs:

$$\hat{y}_T = \frac{f_T(X_T)}{n}$$

where f_T is the regression tree.

Missing values for output variables can be imputed by using the trained RF algorithm on an out-of-training-sample or prediction dataset to compute values based on observed predictors:

$$\hat{y}_P = \frac{f_T(X_P)}{n} \quad (2)$$

RF algorithm has a number of important advantages compared to other data imputation techniques like parametric regression models and non-parameteric k-nearest neighbors machine learning algorithm.¹¹ It allows for using numerical and categorical features, for including non-linear and interaction terms, as well as for handling possible outliers and collinearity among predictors.

Our training dataset is the Bank of Italy's SHIW dataset, which contains self-reported information on the LTV ratios on the mortgages held by the household of the respondents and their income. LTV ratio and mortgagors' income are our output variables, while as features we use the predictors included in the baseline model (immigrant status, loan amount, interest rate type, mortgagors' age, age squared and gender).¹² The prediction dataset is the original CR dataset that does not contain information on LTV and income. We follow the proximity imputation algorithm by using the R package "Random Forest". To achieve a stable error rate, we specified 500 trees.

Regression results are reported in Table 5 for the the samples of single-name and joint mortgages.¹³ In columns [1] and [4], we include LTV ratios with the same controls used in the baseline model of Table 4. Consistent with [Diaz-Serrano and Raya \(2014\)](#), we find a

¹¹For a comparison among data imputation techniques see [Shah et al. \(2014\)](#) and [Tang and Ishwaran \(2017\)](#).

¹²We exclude credit history from our prediction model as SHIW data does not contain this information.

¹³Table A3 in the Appendix show the robustness of results to the inclusion of bank-time and province-time fixed effect and to robust standard errors clustered at the bank level.

positive and statistically significant association between LTV ratios and interest rates paid on mortgages. More importantly, we find that, even after controlling for down payments, the interest rate on single-name (resp., joint) mortgages is for immigrants 20 (resp., 28) bps higher than for native-born Italians.

[TABLE 5 AROUND HERE]

In columns [2] and [5], we add mortgagors' income. As expected, we find that higher earning capacity of mortgagors result in lower mortgage rates. However, even with the same income, immigrants pay more than natives. Finally, in columns [3] and [6] we replicate results classifying mortgagors by the country of origin and, for the sample of joint mortgages and the full sample, distinguishing the loans taken out by two immigrants, two natives and one immigrant and one native co-borrowers. Once again, the findings are consistent with those of the baseline model.

4.4 Oaxaca-Blinder Decomposition

To provide further evidence of unequal treatment between immigrant and native mortgagors, in this section we adopt the Oaxaca-Blinder decomposition approach (Oaxaca, 1973; Blinder, 1973). The aim of the decomposition is to distinguish how much of the average interest rate differential between mortgages made out to immigrants and Italian natives is due to a component that can be explained by differences in explanatory variables assuming that both group of mortgagors receive the same treatment (also referred to as the endowments effect), and an unexplained component arising from differences in the magnitude of the regression coefficients (the coefficient effect). This second component captures the degree of discrimination against immigrant in the mortgage market due to the fact that the same individual characteristic have a different impact on the interest rate of loans granted to immigrant mortgagors and on those granted to native-born Italians.

The Oaxaca-Blinder decomposition of the interest rate gap paid by immigrants starts by running two separate interest rate regressions for two groups of immigrants and native

mortgagors:

$$r_I = X_I' \beta_I + \epsilon_I \quad (3)$$

$$r_N = X_N' \beta_N + \epsilon_N \quad (4)$$

where X' is a vector including the same set of observable characteristics as in column (3) of Table 4. The differential between the mean interest rates paid by immigrants and natives, $\Delta \bar{r} = \bar{r}_I - \bar{r}_N = \bar{X}_I' \hat{\beta}_I - \bar{X}_N' \hat{\beta}_N$, is then decomposed into two components:

$$\Delta \bar{r} = \underbrace{(\bar{X}_I - \bar{X}_N)' \hat{\beta}_N}_{\text{explained component}} + \underbrace{\bar{X}_I' (\hat{\beta}_I - \hat{\beta}_N)}_{\text{unexplained component}} \quad (5)$$

where $\hat{\beta}_N$ is assumed to be the vector of nondiscriminatory coefficients.

Table 6 depicts the results of the decomposition analysis. The interest paid by immigrants on single-name mortgages is 18.4 bps higher than that paid by natives, while the gap is 19.3 bps on joint mortgages.¹⁴

[TABLE 6 AROUND HERE]

The individual characteristics of immigrants account only for 14% of the gap in the mortgage rate. This suggests that if immigrant mortgagors were valued by banks in the same way they value the native ones, the different average characteristics that distinguishes the two groups would justify an interest gap of only 2.64 bps. On the other hand, about 86% of the observed interest gap depends on unexplained factors that are behind the different marginal impact that observable characteristics have for the interest rate charged to immigrant and native mortgagors. The unexplained component is even greater in the case of joint mortgages, representing 93% of the interest gap. Therefore, the Oaxaca-Blinder decomposition confirms the existence of possible different treatments of mortgagors based on national origins.

¹⁴Table A4 in the Appendix reports the underlying regression results of models (3) and (4) for interest rates paid by immigrant and native mortgagors, while in Table A5 we repeat the analysis by using standard errors clustered at the bank level.

4.5 Propensity Score Matching

To identify the effect of the immigrant status of mortgagors on interest rates, we would ideally compare the rates paid by immigrant and native borrowers who were otherwise identical. However, the status of immigrant is not randomly assigned across mortgagors.

To mitigate this concern, we match mortgages held by immigrants to observationally similar mortgages held by Italian natives. We use the propensity score matching (PSM) technique based on a unidimensional balancing score to characterize the probability that the mortgage i is taken up by an immigrant conditional on observables. The objective is to compare mortgages made out to immigrant and native mortgagors that are very similar in terms of the observable characteristics including loan amount, age, credit history, gender and type of loan contract like in our baseline specification. Assuming that there are no significant differences in unobservable characteristics between the two groups of mortgagors, the observed average interest rate gap between mortgages taken up by immigrants and mortgages taken up by the matched group of natives is attributable to the national origin of the mortgagor (Rosenbaum and Rubin, 1983).

We match each immigrant mortgagor to one or more natives on the propensity score by using the nearest neighbor method with replacement, where each native can be paired to more than one immigrant independent of the matching order (Dehejia and Wahba, 2002; Thoemmes and Kim, 2011). In Table 7, we report the average treatment effect on treated (ATT) for matching with one, four and eight mortgages held by Italian natives. Results again indicate that on average immigrants pay significantly more than natives: on single-name mortgages the interest gap is 19 bps, while on the joint mortgages the differential is 33 bps.

[TABLE 7 AROUND HERE]

4.6 Credit history, dominant lenders and cooperative banks

In this section, we try to gauge the role that statistical discrimination and hold-up problems play in mortgage pricing. Precisely, we examine to what extent differences in the availabil-

ity of information about borrowers between banks contribute to explain why banks charge higher interest rates on mortgages taken up by immigrants.

If we suppose that native-born loan officers are more culturally distant and less empathetic towards immigrants than their fellow citizens (Cornell and Welch, 1996; Frame et al., 2022; Jiang et al., 2023), it is reasonable to expect that banks' lack of borrower-specific information and the presence of relationship-specific mortgage search costs for the latter will be more severe for immigrant borrowers rather than for Italian natives. In this case banks have limited information to infer the creditworthiness of individual borrowers and can use observable group characteristics such as race, ethnicity or nationality as a signal of their ability to repay credit, based on the correlation between these group characteristics and risk indicators (such as the delinquency rate). As a result, an immigrant, although otherwise individually identical to a native borrower, is asked for a higher interest rate because she belongs to a different nationality group which is statistically more risky for the bank (Scalera and Zazzaro, 2001). Bank-mortgagor relationships do not typically repeat over time as bank-firm relationships do, and cannot mitigate asymmetric information and statistical discrimination problems (Bellucci et al., 2010; Albareto et al., 2022). However, the amount of information shared in the banking system through the CR narrow the asymmetry of information with potential borrowers for banks and increase competition in the mortgage market, thus reducing statistical discrimination and hold-up problems for mortgagors. We proxy information on a mortgagor available to banks with the number of years from which he/she appears in the Central Credit Register and all the banks in the system can track its credit history. Therefore, we augment the baseline model with an interaction term between the immigrant status and his/her credit history, alternatively measured as a continuous variable or as an indicator variable taking the value of 1 for mortgagors having a credit history longer than 3 years and zero otherwise.

Second, we distinguish between the mortgages granted by the five banks that hold the largest share of the mortgage market (*Main competitors* = 1) and the other fringe competitors (*Main competitors* = 0), and consider the interaction with the immigrant status of the mortgagor. The idea is that the largest competitors rely on automated underwriting systems

for their lending decisions more than small competitors. This can be expected to allow large competitors to charge lower interest rate on average, but at the same time lead them to use statistical information to distinguish among group of borrowers.

Finally, in a similar perspective, we examine whether ethnic disparities in the mortgage interest rates are mitigated for immigrants borrowing from cooperative banks (*Cooperative* = 1). The assumption here is that the cooperative banks have a more in-depth knowledge of the various local economic and social realities in which immigrants are embedded. Furthermore, they are less hierarchically organized, reducing asymmetric information between local loan officers and senior managers at the bank headquarters. For both these reasons, we expect that cooperative banks have more soft information about loan applicants and rely more heavily on soft information rather than on recommendations generated by automated underwriting systems.¹⁵

Results reported in Table 8 suggest that statistical discrimination and hold-up problems play a role in mortgage pricing and explain a part of interest rate differential paid by immigrants. Considering single-name mortgages, specification [1] indicates that immigrants benefit more than natives from the lengthening of their credit history and then the interest gap narrows as their credit history gets longer: for an immigrant with a credit history of 10 years the interest gap reduces by 6.5 bps, while according to column [2] immigrants who entered in Credit Register by more than 3 years pay on average 5.9 bps less than other immigrants (whose interest gap with natives is 25 bps). This suggests that immigrants are ex-ante more opaque than natives and, consequently, gain more from a longer credit history.

Column [3] and [4] show that interest rate disparities against immigrants is on average the same on mortgages granted by dominant and fringe competitors, while immigrants borrowing from cooperative banks pay 10 bps less than immigrants who have the mortgages with other banks.

[TABLE 8 AROUND HERE]

Moving on joint mortgages, results indicate that while mortgagors' credit history (in this

¹⁵Relatedly, a number of recent papers have documented that minority-owned banks and minority loan officers who are culturally closer to minority mortgagors are more likely to approve their applications than other banks or loan officers (Frame et al., 2022; Hurtado and Sakong, 2022; Jiang et al., 2023).

case, the greatest value of *Credit history* between the two co-mortgagors) contributes to reduce, on average, the interest rate paid by borrowers, it does not help to mitigate the interest gap between immigrant and native mortgagors. By contrast, we find evidence that the immigrant interest gap on joint mortgages is greater for the big lenders than for small banks (54 and 32 bps, respectively), although this difference is not statistically significant. Cooperative banks, on the other hand, confirm their ability to accommodate the demand for mortgages from immigrants at prices that are more in line with those charged to Italian natives compared to other banks (precisely, the immigrant interest gap is 23 bps for mortgages with cooperative banks against 43 bps for those with non-cooperative banks).¹⁶

On the whole, these results suggest that the disparity of treatment of immigrant mortgagors that we document in tables 4 and 5, cannot be explained only by a greater difficulty in assessing the individual riskiness of these borrowers accurately, but it also captures some persistent factors such as prejudices arising from cultural distances between lenders and borrowers.

5 Evidence from SHIW data

The loan-level data used so far allowed us to have a very accurate estimate of the parameters of the interest rate model for the variables extracted from the Credit Register. However, the lack of information on the LTV ratio and borrowers' income has forced us to resort to a data imputation process that is subject to possible serious measurement errors.

As an additional robustness check, we run our econometric analysis using survey data from the Bank of Italy's SHIW (Survey on Household Income and Wealth).¹⁷ The SHIW survey is conducted every other year on a representative rotating panel of about 8,000 households resident in Italy. In the SHIW dataset, basic sample unit is the household and the respondent is the head of the household in charge of the household's financial decisions. As

¹⁶We repeat the heterogeneity analysis for the cases of the full sample of single-name and joint mortgages and for the augmented model with imputed LTV ratios and borrowers' income and using joint bank-time and province-time fixed effects and robust standard errors clustered at the bank level. Results, reported in tables from A6 to A9 in the Appendix, are qualitatively and quantitatively almost identical.

¹⁷This is the same survey that we employed above as training set to derive imputed LTV ratios and mortgagors' income.

in the baseline analysis, our key variable of interest is *Immigrant* which takes the value of 1 if the head of the household, defined as the member of the household who is primarily responsible for the financial and economic choices of the household unit, is foreign born. Similarly, self-reported mortgage related information (that is, interest rate, interest rate regime, loan amount, LTV ratio, mortgage destination) are all provided by the head of the household and are asked at the household level. We control for age, education, gender and marital status of the household head and the total income generated by the household. Finally, we include dummies for NUTS2 regions, and dummies for mortgage generating year.

Information on immigrant households are only available since 2006. Therefore, we can use six SHIW waves. For households participating in more than one wave, we include the answers given to the first wave in which information relating to the mortgage taken out by the household appears in the survey and, possibly, the answers relating to a second new mortgage. This leaves us with 2,081 observations. Then, we exclude observations where the reported interest rate is zero or negative or LTV ratio is greater than 1. Our final sample is thus formed by 1,991 mortgages, 91 of which are taken out by immigrants. Table 9 provides definition and summary statistics of variables from the SHIW used in our analysis. Consistent with Credit Register data, the unconditional mean mortgage rate for mortgages taken up by immigrant households is about 33 bps more than that for their native counterparts (3.72% vs 3.39%), even if this gap is not statistically significant.

[TABLE 9 AROUND HERE]

Table 10 presents the regression results using SHIW data. The interest gap on mortgages taken up by immigrants is still statistically significant and it is slightly greater in magnitude (40 bps) compared to the interest gap resulting from the Credit Register data (34 bps). As expected, LTV and fixed rate mortgages exert a positive impact on mortgage rate, while the effect of household income and loan size are negative. Finally, the coefficient of purchase indicates that borrowers pay less on mortgages taken up for purchasing property rather than for renovation.

[TABLE 10 AROUND HERE]

The Bank of Italy monthly provides intermediaries with information on the total debt towards the credit system of each reported customer.

6 Loan approvals

In this last section, we look at loan approvals by examining whether immigrant loan applicants are less likely to obtain mortgages in comparison to native-born Italian applicants.

The Bank of Italy's Central Credit Register shares information among banks in two ways. First, every month banks automatically receive information about firms and households they are currently lending to. In addition, banks can query the CR for information about legal and natural persons they are not lending to (the so-called requests for "CR first information").¹⁸ By law, each bank is allowed to do that only in certain circumstances that are also tracked in the register. One of these occurs when a firm or a natural person asks banks for a loan. In what follows, we focus on these cases only.

Following the approach introduced by Jiménez et al. (2012), we build an indicator for the demand of mortgages at the individual level by using CR data on the banks' requests of first information on natural persons who are not their borrowers. If the bank that has requested information on a particular borrower will grant the loan within a given span of time, say three months, then the loan application is considered successfully approved otherwise it is considered not approved.

In table 11 we report description of variables and descriptive statistics and univariate t-tests for differences between the means of the the groups of Italian native and immigrant loan applicants. Immigrants are 0.8% (resp., 1.66%) less likely to have their loan approved within 3 months (resp., 6 months) from the application than Italian natives.

[TABLE 11 AROUND HERE]

We estimate a linear probability model to test whether immigrants are less able to obtain credit than natives. The equation is the following:

¹⁸Through "CR first information", banks can access information regarding the subject's financial exposure to the banking system, including the number of lending banks, the residual volume of total loans, their technical forms, the use of credit lines and the presence of any defaults or delays in payments.

$$Prob(Approval|_{t-m}d_t = 1)_{ijt} = \alpha \cdot Immigrant_j + \sum_{n=1}^N \beta_n \cdot X_{ijt}^n + \mu_j + \delta_i + \epsilon_{ijt} \quad (6)$$

where, conditional on bank i submitting a request for CR first information on person j within the past m months (that is, conditional on bank i receiving a demand for a mortgage by j , $_{t-m}d_t = 1$), *Approval* is a dummy that takes the value 1 if bank i grants a loan to person j at time t , and zero otherwise.

Our sample includes more than 7 million observations related to all information requests on natural persons posted by Italian banks in 2015 and 2016. Regression results are reported in Table 12. We use the 3-months period to assess whether a loan has been approved, conditionally on a loan request, while results for the 6-months period are reported in the Appendix, table A10. Our baseline specifications in odd columns control for the age (and age squared) and the gender of the prospective mortgagor, her credit history in the CR and the presence of co-mortgagors. Then, in even columns we estimate specifications including the imputed mortgagors' income derived using machine learning RF procedure as in the loan pricing model in section 4.3. In this case, however, the training dataset employed to elicit the imputed income includes all the respondent households in the six SHIW waves from 2006 to 2016. Finally, all specifications include bank and province fixed effects and use robust standard errors.¹⁹

[TABLE 12 AROUND HERE]

Consistent with the hypothesis of disparities in access to credit, immigrants have a lower probability of seeing their loan applications approved by banks. The coefficient estimated for *Immigrant* is negative and statistically significant in both specifications in columns [1] and [2], indicating that immigrants are on average 2,7% less successful than natives when they ask banks for a mortgage. All control variables have the expected sign, including imputed income, reassuring us of the reliability of our results.

In columns [3]-[8], we augment the model with interactions for borrowers' credit history and distinguish mortgages provided by dominant competitors and cooperative banks.

¹⁹Again, results are robust to using joint time-bank and time-province fixed effects and to standard errors clustered at the bank level (see the Appendix, tables A11 and A12).

As for the pricing model, disparities in access to credit for immigrants are statistically and economically significantly lower when they apply to cooperative banks: in this case, the approval gap with Italian natives is only 1.5%. This confirms the positive contribution of local, community banks to social integration of immigrants. By contrast, the approval gap widens when immigrants direct their loan applications to dominant competitors, becoming equal to 3.2%. Finally, credit history does not help to mitigate disparity in access to credit, leaving the approval gap economically unaffected.

7 Conclusions

In this paper we have analyzed empirically the interest rate formation in the Italian mortgage market and the probability of mortgage approvals by Italian banks, looking at a possible discrimination against immigrant borrowers. In spite of the importance of affordable access to mortgages for homeownership and social integration of immigrants and ethnic minorities in European countries (Arbaci and Malheiros, 2010; Accetturo et al., 2014; Kara and Molyneux, 2017), this issue is still largely unexplored in the literature.²⁰

We have conducted a number of empirical analyses using loan data from the Italian Credit Register for the period 2011-2016 and survey data from the Bank of Italy's Survey on Household Income and Wealth (SHIW) for the period 2006-2011. Our findings consistently document that on average the interest rate on mortgages taken up by immigrants is significantly higher than the interest rate paid by Italian natives. We analyze both single-name and joint mortgage contracts: on the latter, immigrants pay 42 basis points more than Italians, while on single-name mortgages the interest gap is 25 basis points. The interest rate differential is highest for Asian, Central American and African immigrants and it remains statistically significant even after controlling for many loan and individual characteristics. In particular, since Credit Register does not include information on mortgagors' income and mortgages' LTV ratio, first we used random forest machine learning technique to obtain imputed values for these two important interest rate determinants from information contained

²⁰o the best of our knowledge, the only exception is Diaz-Serrano and Raya (2014) who analyze the case of Spain.

in the SHIW and Wealth, and second we repeated our analysis on SHIW data. In both cases, we confirmed our baseline results. Oaxaca-Blinder decomposition indicates that the greatest part of the interest rate differential (85% and 93% for single-name and joint mortgages, respectively) can be imputed to the unexplained component capturing possible discrimination effects. Furthermore, we showed that the interest gap is lower for immigrants with a long credit history but it remains statistically significant, thus suggesting that statistical discrimination cannot explain the gap entirely.

As a final step, we have examined the probability of mortgage approvals for the the sub-period 2015-2016. Once again, our findings indicate that for immigrants the probability of having their mortgage application approved is on average statistically lower than for natives.

References

- Accetturo, A., G. Barboni, M. Cascarano, and E. Garcia-Appendini (2023). Propensity score-matching methods for nonexperimental causal studies. *Journal of Financial Intermediation* 53(Article No. 101018).
- Accetturo, A., F. Manaresi, S. Mocetti, and E. Olivieri (2014). Don't stand so close to me: The urban impact of immigration. *Regional Science and Urban Economics* 45, 45–56.
- Albareto, G., P. E. Galardo, Maddalena Mistrulli, and B. Sorvillo (2022). Culture, lending relationships, and the cost of credit. *Review of Corporate Finance Studies* 11(3), 736–774.
- Alesina, A. F., F. Lotti, and P. E. Mistrulli (2013). Do women pay more for credit? Evidence from Italy. *Journal of the European Economic Association* 11(s1), 45–66.
- Ambrose, B. W. J., J. N. Conklin, and L. A. Lopez (2021). Does borrower and broker race affect the cost of mortgage credit? *Review of Financial Studies* 34(2), 790–826.
- Arbaci, S. and J. Malheiros (2010). De-segregation, peripheralisation and the social exclusion of immigrants: Southern european cities in the 1990s. *Journal of Ethnic and Migration Studies* 36(2), 227–255.
- Arrow, K. (1973). The theory of discrimination. In O. Ashenfelter and A. Rees (Eds.), *Discrimination in labor markets*, pp. 3–33. Princeton, NJ: Princeton University Press.
- Bank of Italy (2021). Rilevazione dei dati granulari sul credito: istruzioni per gli intermediari segnalanti. Circolare n. 297 del 16 maggio 2017 - 3° aggiornamento del dicembre 2021, Bank of Italy, Rome: Italy.
- Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics* 143(1), 30–56.
- Bayer, P., F. Ferreira, and S. L. Ross (2014). Race, ethnicity and high-cost mortgage lending. NBER Working Paper Series 20762, National Bureau of Economic Research.

- Bayer, P., F. Ferreira, and S. L. Ross (2017). What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders. *Review of Financial Studies* 31(1), 175–205.
- Becker, G. (1971). *The Economics of Discrimination* (2nd ed.). Chicago, ILL: University of Chicago Press.
- Bellucci, A., A. Borisov, and A. Zazzaro (2010). Does gender matter in bank–firm relationships? Evidence from small business lending. *Journal of Banking and Finance* 34(12), 2968–2984.
- Bertocchi, G., M. Brunetti, and A. Zaiceva (2022). The financial decisions of immigrant and native households: Evidence from Italy. *Italian Economic Journal* (in press).
- Bhutta, N. and A. Hizmo (2021). Do minorities pay more for mortgages? *Review of Financial Studies* 34(2), 763–789.
- Black, A., T. P. Boehm, and R. P. DeGennaro (2003). Is there discrimination in mortgage pricing? The case of overages. *Journal of Banking and Finance* 27(6), 1139–1165.
- Black, H., R. L. Schweitzer, and L. Mandell (1978). Discrimination in mortgage lending. *American Economic Review* 68(2), 186–191.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4), 436–455.
- Bocian, D. G., K. S. Ernst, and W. Li (2008). Race, ethnicity and subprime home loan pricing. *Journal of Economics & Business* 60(1), 110–124.
- Boehm, T. P. and A. Schlottmann (2007). Mortgage pricing differentials across Hispanic, African-American, and white households: Evidence from the American housing survey. *Cityscape* 9(2), 93–136.
- Boehm, T. P., P. D. Thistle, and A. Schlottmann (2006). Rates and race: An analysis of racial disparities in mortgage rates. *Housing Policy Debate* 17(1), 109–149.

- Borjas, G. (2002). Homeownership in the immigrant population. *Journal of Urban Economics* 52(3), 448–476.
- Bostic, R. W. (2003). A test of cultural affinity in home mortgage lending. *Journal of Financial Services Research* 23(1), 89–112.
- Breiman, L. (1996). Bagging predictors. *Machine learning* 24(2), 123–140.
- Breiman, L. (2001). Random forests. *Machine learning* 45(1), 5–32.
- Calcagnini, G., G. Giombini, and E. Lenti (2015). Gender differences in bank loan access: An empirical analysis. *Italian Economic Journal* 1(2), 193–217.
- Cheng, P., Z. Lin, and Y. Liu (2015). Racial discrepancy in mortgage interest rates. *Journal of Real Estate Finance and Economics* 51(1), 101–120.
- Chiuri, M. C. and T. Jappelli (2003). Financial market imperfections and home ownership: A comparative study. *European Economic Review* 47(5), 857–875.
- Constant, A., R. Roberts, and K. Zimmermann (2009). A model of housing tenure choice. *Urban Studies* 46(9), 1879–1898.
- Cornell, B. and I. Welch (1996). Culture, information, and screening discrimination. *Journal of Political Economy* 104(3), 542–571.
- Courchane, M. and D. Nickerson (1997). Discrimination resulting from overage practices. *Journal of Financial Services Research* 11(1), 133–151.
- Crawford, G. W. and E. Rosenblatt (1999). Differences in the cost of mortgage credit implications for discrimination. *Journal of Real Estate Finance and Economics* 19(2), 147–159.
- Dehejia, R. H. and S. Wahba (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics* 84(1), 151–161.
- Delis, M. D. and P. Papadopoulos (2019). Mortgage lending discrimination across the U.S.: New methodology and new evidence. *Journal of Financial Services Research* 56(3), 341–368.

- Diaz-Serrano, L. and J. M. Raya (2014). Mortgages, immigrants and discrimination: An analysis of the interest rates in Spain. *Regional Science and Urban Economics* 45(1), 22–32.
- Elmer, P. J. and S. A. Seelig (1999). Insolvency, trigger events, and consumer risk posture in the theory of single-family mortgage default. *Journal of Housing Research* 10(1), 1–25.
- Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010). What “triggers” mortgage default? *American Economic Review* 100(2), 490–494.
- EMF-ECBC (2021). Hypostat 2021. A review of Europe’s mortgage and housing markets. Technical report, European Mortgage Federation - European Covered Bond Council, Bruxelles.
- Fisman, R., D. Paravisini, and V. Vig (2017). Cultural proximity and loan outcomes. *American Economic Review* 107(2), 457–92.
- Frame, W. S., R. Huang, E. J. Mayer, and A. Sunderam (2022). The impact of minority representation at mortgage lenders. NBER Working Paper Series 30125, National Bureau of Economic Research.
- Ghent, A. C., R. Hernandez-Murillo, and O. M.T. (2014). Differences in subprime loan pricing across races and neighborhoods. *Regional Science and Urban Economics* 48, 199–215.
- Goodman, L., A. McCargo, E. Golding, B. Bai, and S. Storchak (2018). Barriers to accessing homeownership. down payment, credit, and affordability. Technical report, Housing Finance Policy Center.
- Guiso, L. and T. Jappelli (2002). Private transfers, borrowing constraints and the timing of homeownership. *Journal of Money, Credit and Banking* 34(2), 315–339.
- Guiso, L., A. Pozzi, A. Tsoy, L. Gambacorta, and P. E. Mistrulli (2022). The cost of steering in financial markets: Evidence from the mortgage market. *Journal of Financial Economics* 143(3), 1209–1226.
- Hanson, A., Z. Hawley, H. Martin, and B. Liu (2016). Discrimination in mortgage lending: Evidence from a correspondence experiment. *Journal of Urban Economics* 92, 48–65.

- Haurin, D., P. Hendershott, and S. Wachter (1997). Borrowing constraints and the tenure choice of young households. *Journal of Housing Research* 8(2), 137–154.
- Henderson, J. and Y. Ioannides (1983). A model of housing tenure choice. *American Economic Review* 73(1), 98–113.
- Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Volume 1, pp. 278–282. IEEE.
- Hurtado, A. and J. Sakong (2022). The effect of minority bank ownership on minority credit. Working Paper Series 325, Federal Reserve Bank of Chicago.
- Jiang, E. X., Y. Lee, and W. S. Liu (2023). Reducing racial disparities in consumer credit: the role of minority loan officers in the era of algorithmic underwriting. Research paper series, USC Marshall School of Business.
- Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina (2012). Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *American Economic Review* 102(5), 2301–26.
- Kara, A. and P. Molyneux (2017). Household access to mortgages in the UK. *Journal of Financial Services Research* 52(3), 253–275.
- Kim, S. and G. Squires (1998). The color of money and the people who lend it. *Journal of Housing Research* 9(2), 271–284.
- LaCour-Little, M. (1999). Discrimination in mortgage lending: A critical review of the literature. *Journal of Real Estate Literature* 7(1), 15–49.
- Ladd, H. F. (1998). Evidence on discrimination in mortgage lending. *Journal of Economic Perspectives* 12(2), 41–62.
- Lang, J. H., M. Pirovano, M. Rusnák, and C. Schwarz (2020). Mortgage markets. *Financial Stability Review*, European Central Bank.

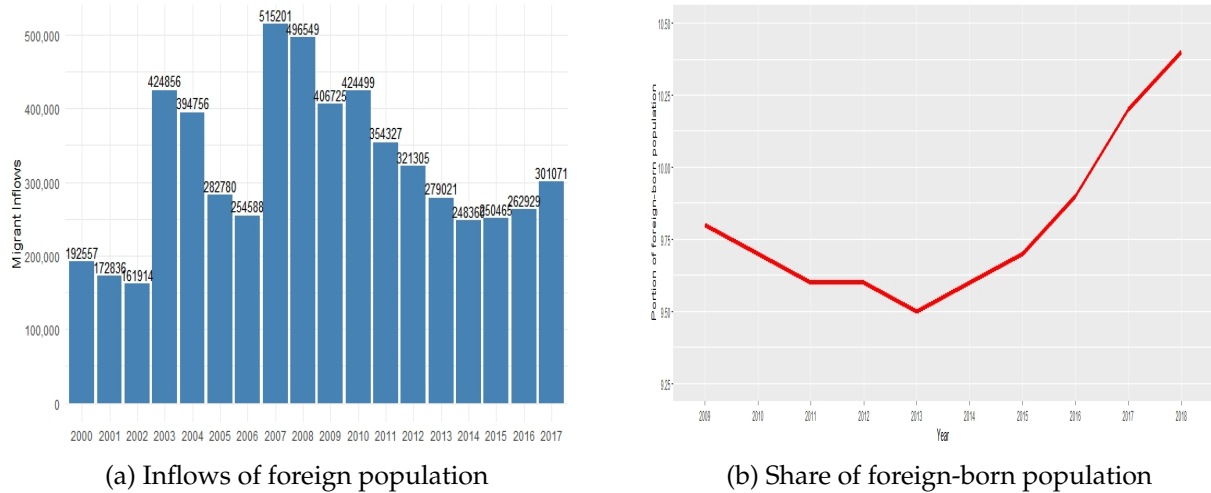
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14(3), 693–709.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review* 62(4), 659–661.
- Reid, C., D. Bocian, W. Li, and R. Quercia (2017). Revisiting the subprime crisis: The dual mortgage market and mortgage defaults by race and ethnicity. *Journal of Urban Affairs* 39(4), 469–487.
- Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), 41–55.
- Ross, S. L., M. A. Turner, E. Godfrey, and R. R. Smith (2008). Mortgage lending in Chicago and Los Angeles: A paired testing study of the preapplication process. *Journal of Urban Economics* 63(3), 902–919.
- Scalera, D. and A. Zazzaro (2001). Group reputation and persistent (or permanent) discrimination in credit markets. *Journal of Multinational Financial Management* 11(4-5), 483–496.
- Shah, A. D., J. W. Bartlett, J. Carpenter, O. Nicholas, and H. Hemingway (2014). Comparison of random forest and parametric imputation models for imputing missing data using MICE: A CALIBER study. *American Journal of Epidemiology* 179(6), 764–774.
- Stefani, M. L. and V. Vacca (2013). Credit access for female firms: evidence from a survey on European SMEs. Occasional papers 176, Bank of Italy.
- Tang, F. and H. Ishwaran (2017). Random forest missing data algorithms. *Statistical Analysis and Data Mining* 10(6), 363–377.
- Thoemmes, F. J. and E. S. Kim (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research* 46(1), 90–118.
- Trucchi, S. (2016). Credit markets and housing choices. *Bulletin of Economic Research* 68(S1), 1–19.

Turner, M. A. and F. Skidmore (1999). *Mortgage lending discrimination: A review of existing evidence*. Washington DC: The Urban Institute.

Wei, B. and F. Zhao (2022). Racial disparities in mortgage lending: New evidence based on processing time. *Review of Corporate Finance Studies* (in press).

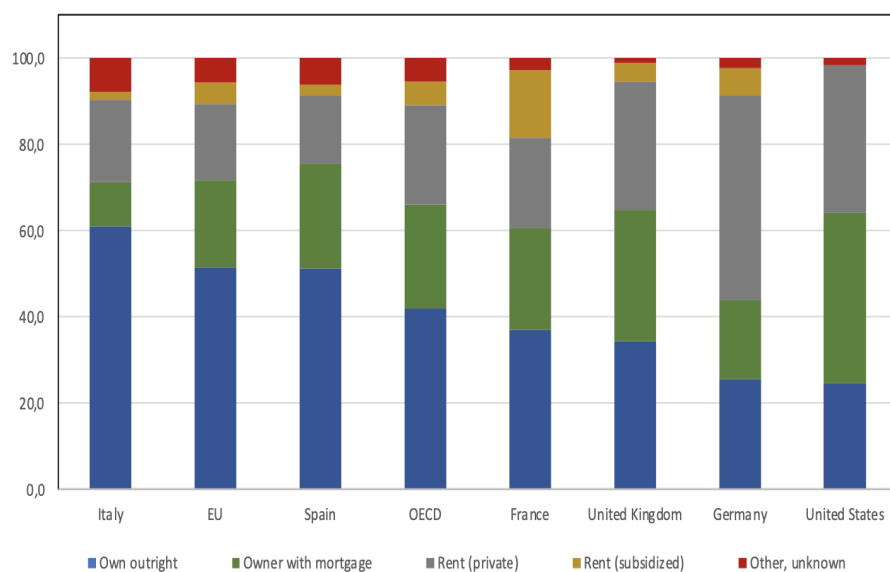
Main figures and tables

Figure 1: Foreign population in Italy.



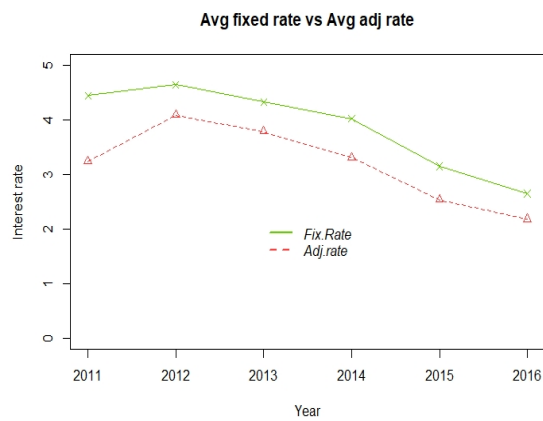
Note. Plot (a) shows the number of foreign-born people legally entered in Italy by year; plot (b) shows the stock of foreign-born population in Italy as percentage of total population. Data source: the "OECD International Migration Database" available at <https://www.oecd.org/migration/keystat.htm>.

Figure 2: Housing occupancy rates.

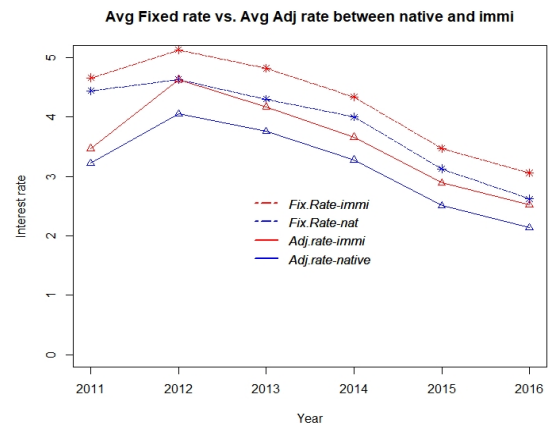


Note. This Figure shows the percentage of households in different tenure types in latest year available (2019 or 2018). Data source: OECD, Housing Market database, available at <https://www.oecd.org/housing/data/affordable-housing-database/housing-market.html>.

Figure 3: Interest rate on mortgage loans.



(a) ARMs versus FRMs



(b) ARMs versus FRMs for immigrants and native

Note. Plot (a) shows the average interest rate in percentage on adjusted-rate and fixed-rate mortgage contracts; plot (b) shows the average interest rates (%) paid by immigrant and native mortgagors on ARM and FRM contracts. Data source: Bank of Italy, Central Credit Register.

Table 1: Top 20 countries of origin of immigrants residing in Italy.

2011		2016	
Romania	9,88,576	Romania	1,151,395
Albania	482,627	Albania	467,687
Morocco	452,424	Morocco	437,485
China	209,934	China	271,330
Ukraine	200,730	Ukraine	230,728
Philippines	134,154	Philippines	165,900
Moldova	130,948	India	150,456
India	121,036	Moldova	142,266
Poland	109,018	Bangladesh	118,790
Tunisia	106,291	Egypt	109,871
Perù	98,603	Perù	103,714
Ecuador	91,625	Sri Lanka	102,316
Egypt	90,365	Pakistan	101,784
Macedonia	89,900	Senegal	98,176
Bangladesh	82,451	Poland	97,986
Sri Lanka	81,094	Tunisia	95,645
Senegal	80,989	Ecuador	87,427
Pakistan	75,720	Nigeria	77,264
Nigeria	53,613	Macedonia	73,512
Serbia	52,954	Bulgaria	58,001

Note. Foreign-born population by country of origin. Data source: the Italian National Statistics Institute (ISTAT), available at <http://dati.istat.it>.

Table 2: Variable description and statistics: Mortgage pricing analysis.

Variable	Description	N	Mean (SD)
<i><u>Dependent variable</u></i>			
APR	Interest rate on mortgage loans to individual i by bank j (in percentage).	954,341	3.22 (1.15)
<i><u>Mortgage characteristics</u></i>			
Loan amount	Amount of outstanding loans (in log).	954,341	11.74 (0.45)
Fixed	Dummy variable, equal to 1 if a loan is fixed rate, and 0 otherwise.	954,341	0.28 (0.45)
Coint	Dummy variable, equal to 1 if there is a co-borrower in the loan contract, and 0 otherwise.	954,341	0.48 (0.49)
Singleimmi	Dummy variable, equal to 1 if the contract has only one immigrant borrower, and 0 otherwise	954,341	0.03 (0.15)
Cointimmi	Dummy variable, equal to 1 if in a contract both primary and Co-borrower are immigrant, and 0 otherwise	954,341	0.03 (0.17)
Cointnat	Dummy variable, equal to 1 if in a contract both primary and Co-borrower are native, and 0 otherwise.	954,341	0.42 (0.49)
Cointmix	Dummy variable, equal to 1 if in a contract immigrant and native participate together as primary and Co-borrower, and 0 otherwise.	954,341	0.03 (0.17)
<i><u>Borrower characteristics</u></i>			
Immigrant	Dummy variable, equal to 1 if the borrower is non-Italian, and 0 otherwise.	954,341	0.06 (0.23)
Europa	Dummy variable, equal to 1 if the borrower is European except Italian, and 0 otherwise.	954,341	0.04 (0.19)
Asia	Dummy variable, equal to 1 if the borrower is Asian, and 0 otherwise.	954,341	0.006 (0.07)
Africa	Dummy variable, equal to 1 if the borrower is African, and 0 otherwise.	954,341	0.004 (0.06)
North America	Dummy variable, equal to 1 if the borrower is North American, and 0 otherwise.	954,341	0.0008 (0.02)

Continued on next page

Table 2: Variable description and statistics: Mortgage pricing analysis – continued from previous page

Variable	Description	N	Mean (SD)
Central America	Dummy variable, equal to 1 if the borrower is Central American, and 0 otherwise.	954,341	0.0005 (0.02)
South America	Dummy variable, equal to 1 if the borrower is South American, and 0 otherwise.	954,341	0.005 (0.07)
Oceania	Dummy variable, equal to 1 if the borrower is Oceanian, and 0 otherwise.	954,341	0.0002 (0.014)
Age	Primary borrower age in years.	954,341	39.79 (9.73)
Male	Dummy variable, equal to 1 if the borrower is male, and 0 otherwise.	954,341	0.56 (0.49)
<i>Moderator factors</i>			
Main competitors	Dummy variable, equal to 1 for the 5 banks holding the largest share of mortgages in our sample, which amount to 39%, and 0 otherwise.	954,341	0.39 (0.48)
Cooperative	Dummy variable, equal to 1 if the bank is a co-operative bank, which comprises 13.39% of the sample, and 0 otherwise.	954,341	0.1339 (0.34)
Credit history	Number of years since first time appearance in the Central Credit Register.	954,341	2.23 (4.19)

Notes. Data source: Bank of Italy, Central Credit Register and Analytical Survey on Loan Interest Rates, 2011-2016.

Table 3: Univariate analysis.

	Italian natives	Immigrants	Diff.	t-stat
<u>Panel A: Single-name mortgages</u>				
APR	3.16 (1.16)	3.34 (1.09)	-0.18***	-25.66
Loan amount (ln)	11.70 (0.45)	11.65 (0.43)	0.05***	17.81
Fixed	0.27 (0.45)	0.25 (0.43)	0.02***	8.21
Age (years)	39.98 (9.60)	39.78 (8.90)	0.20***	3.44
Male	0.61 (0.49)	0.53 (0.50)	0.08***	24.79
Credit history	2.74 (4.36)	2.06 (3.79)	0.68***	27.47
Observations	465,450	24,739		
<u>Panel B: Joint mortgages</u>				
APR	3.26 (1.15)	3.58 (1.07)	-0.31***	-49.71
Loan amount (ln)	11.79 (0.43)	11.66 (0.32)	0.13***	70.55
Fixed	0.30 (0.46)	0.26 (0.44)	0.04***	16.78
Age (years)	39.74 (10.07)	37.46 (8.10)	2.30***	47.45
Male	0.51(0.50)	0.51(0.50)	-0.00***	-0.65
Credit history	1.78 (4.01)	0.40 (1.88)	1.39***	113.16
Observations	433,063	31,089		
<u>Panel C: Full sample</u>				
APR	3.20(1.15)	3.47(1.09)	-0.26***	-55.53
Loan amount (ln)	11.74 (0.44)	11.66 (0.38)	0.09***	54.59
Fixed	0.29 (0.45)	0.25 (0.43)	0.03***	17.09
Age (years)	39.87 (9.82)	38.49 (8.54)	1.38***	36.82
Male	0.56(0.50)	0.52(0.50)	0.04***	19.61
Credit history	2.30 (4.23)	1.14 (3.00)	1.15***	85.47
Observations	898,513	55,828		

Notes. This tables reports summary statistics for mortgages made out to Italian natives and immigrants, and t-statistics for differences in mean. Single-name mortgages include loans made out to a single borrower. Joint mortgages include loans to two co-borrowers, either both Italian natives or both immigrants, excluding the "mixed" joint contracts made out to a native and an immigrant borrower. Full sample consists of all single-name and joint mortgages. Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 4: Mortgage pricing regression results: baseline model.

	Single-name mortgages			Joint mortgages		Full sample
	[1]	[2]	[3]	[4]	[5]	[6]
Immigrant	0.247*** (0.00608)	0.236*** (0.00563)				
ln(Loan amount)		-0.270*** (0.00329)	-0.270*** (0.00329)		-0.253*** (0.00363)	-0.262*** (0.00243)
Fixed		0.747*** (0.00362)	0.747*** (0.00362)		0.761*** (0.00348)	0.751*** (0.00251)
Age		-0.00826*** (0.000951)	-0.00818*** (0.000951)		-0.0208*** (0.000895)	-0.0147*** (0.000651)
Age-squared		9.19e-05*** (1.11e-05)	9.12e-05*** (1.11e-05)		0.000192*** (1.03e-05)	0.000142*** (7.54e-06)
Male		0.0226*** (0.00274)	0.0222*** (0.00274)		0.0110*** (0.00262)	0.0126*** (0.00188)
Credit history		-0.0134*** (0.000352)	-0.0134*** (0.000352)		-0.00686*** (0.000376)	-0.00965*** (0.000255)
Europa			0.234*** (0.00685)			
Asia			0.318*** (0.0165)			
Africa			0.248*** (0.0191)			
North America			0.0701** (0.0316)			
Central America			0.262*** (0.0486)			
South America			0.216*** (0.0168)			
Oceania			0.0722 (0.0577)			
Cointimmi				0.428*** (0.00534)	0.404*** (0.00499)	0.489*** (0.00490)
Cointmix					0.117*** (0.00495)	0.208*** (0.00495)
Singleimmi						0.246*** (0.00558)
Cointnat						0.0950*** (0.00202)
Time FE	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
R-squared	0.285	0.367	0.367	0.314	0.400	0.383
Observations	490,189	490,189	490,189	464,152	464,152	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. Robust standard errors are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table 5: Mortgage pricing regression results: augmented model with imputed LTV ratio and mortgagors' income.

	Single-name mortgages			Joint mortgages		Full sample
	[1]	[2]	[3]	[4]	[5]	[6]
Immigrant	0.218*** (0.00682)	0.195*** (0.00681)				
ln (Loan amount)	-0.287*** (0.00462)	-0.250*** (0.00472)	-0.250*** (0.00472)	-0.343*** (0.00494)	-0.311*** (0.00499)	-0.279*** (0.00337)
Fixed	0.761*** (0.00359)	0.751*** (0.00360)	0.751*** (0.00360)	0.799*** (0.00344)	0.773*** (0.00349)	0.775*** (0.00251)
Age	-0.00882*** (0.000944)	0.00978*** (0.00106)	0.00984*** (0.00106)	-0.0239*** (0.000876)	-0.00427*** (0.000957)	-0.000305 (0.000689)
Age-squared	0.000102*** (1.11e-05)	-7.54e-05*** (1.20e-05)	-7.61e-05*** (1.20e-05)	0.000246*** (1.02e-05)	6.04e-05*** (1.09e-05)	1.82e-05** (7.87e-06)
Male	0.0236*** (0.00273)	0.104*** (0.00328)	0.104*** (0.00328)	0.0101*** (0.00253)	0.0147*** (0.00252)	0.0514*** (0.00189)
Credit history	-0.0134*** (0.000349)	-0.0132*** (0.000349)	-0.0132*** (0.000349)	-0.00614*** (0.000368)	-0.00576*** (0.000368)	-0.00857*** (0.000249)
LTV	0.00155*** (0.000299)	0.00113*** (0.000299)	0.00109*** (0.000299)	0.00801*** (0.000312)	0.00737*** (0.000312)	0.00376*** (0.000213)
ln (Income)		-0.365*** (0.00914)	-0.365*** (0.00914)		-0.384*** (0.00691)	-0.314*** (0.00526)
Europa			0.193*** (0.00785)			
Asia			0.279*** (0.0167)			
Africa			0.205*** (0.0192)			
North America			0.0242 (0.0314)			
Central America			0.221*** (0.0472)			
South America			0.172*** (0.0170)			
Oceania			0.0427 (0.0567)			
Cointimmi				0.313*** (0.00612)	0.284*** (0.00611)	0.384*** (0.00553)
Cointmix				0.118*** (0.00475)	0.116*** (0.00473)	0.167*** (0.00479)
Singleimmi						0.176*** (0.00601)
Cointnat						0.0543*** (0.00208)
Time FE	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES
R-squared	0.377	0.379	0.379	0.441	0.445	0.424
Observations	490,189	490,189	490,189	464,152	464,152	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. LTV is the ratio between the amount of the mortgage and the lender self-assessed value of the house. Income is the mortgagors' income. Both LTV and Income are imputed variables obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Robust standard errors are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table 6: Oaxaca-Blinder decomposition of mortgage interest rate gap between immigrants and Italian natives

	Single name mortgages		Joint mortgages		Full sample	
	Mean interest rate	Decomposition	Mean interest rate	Decomposition	Mean interest rate	Decomposition
Immigrants	3.340*** (0.00696)		3.579*** (0.00606)		3.474*** (0.00460)	
Natives	3.157*** (0.00170)		3.261*** (0.00181)		3.205*** (0.00124)	
Differential	0.184*** (0.00716)		0.319*** (0.00633)		0.269*** (0.00476)	
Explained components		0.0269*** (0.00160)		0.0479*** (0.00141)		0.0443*** (0.000992)
Unexplained components		0.157*** (0.00701)		0.271*** (0.00633)		0.225*** (0.00470)
Observations	490,189	490,189	432,191	432,191	922,380	922,380

Notes. This table shows results from Oaxaca-Blinder decomposition of the mean interest rate differential between mortgages made out to immigrants and native-born Italians. The explained component (or endowments effect) is the part of the interest rate differential that is attributable to group differences in explanatory variables. The unexplained component (or coefficients effect) is the part of the interest rate differential that is attributable to differences in coefficients of explanatory variables. The decompositions are conducted by excluding all samples where immigrants and natives participate jointly in a contract, in order to find a clear distinction between the differences in mortgage rates between natives and immigrants. For joint mortgages, immigrant refers to the contracts where both participants are immigrants, while in the full sample, immigrant refers to contracts where either one or both participants are immigrants. The underlying regressions on interest rates charged on mortgages to immigrants and natives are reported in Table A4 in the Appendix. A supplementary table in the appendix depicts the results of the Oaxaca-Blinder decomposition using cluster standard errors (see table A5). Robust standard errors in parentheses. *** p -value < 0.01, ** p -value < 0.05, * p -value < 0.1.

Table 7: The matching procedure: Average treatment effect on treated

No. of controls matched	Single-name mortgages	Joint mortgages	Full sample
	Average treatment effect (ATT)	Average treatment effect (ATT)	Average treatment effect (ATT)
n=1	0.192*** (0.00862)	0.331*** (0.0099)	0.256*** (0.006)
n=4	0.189*** (0.00717)	0.3305*** (0.0085)	0.259*** (0.0052)
n=8	0.1904*** (0.0069)	0.3312*** (0.0082)	0.2579*** (0.005)

Notes. This table reports the difference in mean between mortgage interest rates paid by immigrants (treated units) and natives nearest to immigrants (non-treated units). Non-treated mortgages are assigned to the control group base on propensity score estimates for the probability of a loan is taken up by immigrants. We ran the "teffects" command in Stata to estimate the average treatment effect (ATT) using a nearest-neighbor matching method with a caliper of 0.01 without replacement on the sample of mortgages to natural persons included in the Bank of Italy' Central Credit Register in the period 2011-2016.

*** p -value < 0.01, ** p -value < 0.05, * p -value < 0.1.

Table 8: Mortgage pricing regression results: borrowers' credit history, dominant lenders and cooperative banks

	Single-name mortgages				Joint mortgages			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Immigrant	0.2498*** (0.0063)	0.2514*** (0.0063)	0.2348*** (0.0067)	0.2512*** (0.0062)				
Cointimmi					0.4043*** (0.0051)	0.4042*** (0.0051)	0.3243*** (0.0060)	0.4319*** (0.0055)
Credit history	-0.0132*** (0.0004)				-0.0068*** (0.0004)			
Immigrant $\times$ Credit history	-0.0065*** (0.0016)							
Cointimmi $\times$ Credit history					-0.0018 (0.0024)			
$\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.1245*** (0.0032)				-0.0672*** (0.0038)		
Immigrant $\times \mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.0588*** (0.0138)						
Cointimmi $\times \mathbb{1}(\text{Credit history} > 3 \text{ yrs})$						-0.0158 (0.0224)		
Immigrant $\times$ Main competitors			0.0045 (0.0123)					
Cointimmi $\times$ Main competitors							0.2125*** (0.0103)	
Immigrant $\times$ Cooperative				-0.1024*** (0.0148)				
Cointimmi $\times$ Cooperative								-0.1997*** (0.0122)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov FE	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.3672	0.3673	0.3672	0.3672	0.4003	0.4002	0.4008	0.4005
Observations	490,189	490,189	490,189	490,189	464,152	464,152	464,152	464,152

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of moderators is reported in table 2. All columns include the same control variables of specifications [2] and [5] in Table 4. Table A7, a supplementary table in the appendix reports the coefficients from OLS regressions using the same specifications with Bank $\times$ Time and Province $\times$ Time joint fixed effects, and clustered standard errors. Robust standard errors are in parentheses. *** p -value < 0.001 , ** p -value < 0.01 , * p -value < 0.05 .

Table 9: Variables from SHIW surveys. Description and statistics

Variable	Description	N	Native Mean (S.D)	Immigrant Mean (S.D)	Diff.	t-stat
<i>Dependent variable</i>						
Mortgage Rate	Interest rate charged on mortgage loans to household i (in percentage)	1,955	3.35 (2.35)	3.59 (2.22)	-0.24	-1.01
<i>Mortgage Characteristics</i>						
Loan amount	The initial amount of mortgage loan (in log).	1,531	11.22 (0.80)	11.55 (0.45)	-0.33***	-5.76
Fixed	Dummy variable, equal to 1 if the loan is fixed rate, and 0 otherwise.	1,955	0.41 (0.49)	0.34 (0.47)	0.07	1.35
LTV	Loan amount over the house value.	1955	60.98 (29.01)	86.73 (19.57)	-25.75***	-11.87
Purchase	Dummy variable, equal to 1 if a mortgage loan is for the purchase of household residence, and 0 otherwise.	1,955	0.67 (0.47)	0.77 (0.42)	-0.10**	-2.18
Duration	Number of years to repay the mortgage.	1,955	18.10 (7.62)	22.54 (6.39)	-4.45***	-6.39
<i>Borrower characteristics</i>						
Immigrant	Dummy variable, equal to 1 if the head of the household is foreign-born, and 0 otherwise.	1,955	1,865	90		
Income	The disposable income of the head of the household (in log).	1,948	10.02 (0.72)	9.5 (0.89)	0.53***	5.93
Age	Age of the head of the household in years.	1,955	42.05 (11.61)	36.12 (8.17)	5.93***	6.22
Male	Dummy variable, equal to 1 if the head of the household is male, and 0 otherwise.	1,263	0.65 (0.48)	0.66 (0.48)	-0.01	-0.19
Education	Categorical variable (1-8) that indicates the educational level of the head of the household; it takes value 1 for the head of the household having no education and value 8 for postgraduate qualification	1,955	4.60 (1.57)	4.36 (1.46)	0.24	1.52
Married	Dummy variable, equal to 1 if the head of household is married and 0 otherwise.	1955	0.75 (0.43)	0.86 (0.33)	0.10**	-2.66

Notes. Data source: Bank of Italy, The Survey on Household Income and Wealth (SHIW), 2006-2016. This table presents the description of variables used in the regression analysis of Table 10, summary statistics for immigrant and Italian natives household heads, and t-statistics for differences in mean.

Table 10: Regression results for SHIW data

	[1]	[2]	[3]
Immigrant	0.4795*** (0.1575)	0.4771*** (0.1589)	0.3988** (0.1666)
Purchase	-0.8023*** (0.0913)	-0.7873*** (0.0945)	-0.7798*** (0.0948)
LTV	0.0054*** (0.0018)	0.0054*** (0.0018)	0.0040** (0.0019)
Fixed	0.6376*** (0.0821)	0.6358*** (0.0822)	0.6315*** (0.0821)
ln(Loan size)	-0.4785*** (0.0775)	-0.4454*** (0.0842)	-0.3618*** (0.0902)
Duration		-0.0053 (0.0068)	-0.0085 (0.0071)
Age		-0.0276 (0.0211)	-0.0236 (0.0213)
Age-squared		0.0003 (0.0002)	0.0002 (0.0002)
Male		0.0363 (0.0848)	0.1081 (0.0936)
Education			-0.0457 (0.0305)
Married			0.0024 (0.0998)
ln(Income)			-0.1437** (0.0656)
Time FE	YES	YES	YES
Region FE	YES	YES	YES
R-squared	0.2709	0.2720	0.2766
Observations	1,524	1,524	1,518

Notes. This table presents the coefficients from OLS regressions. The dataset includes self-reported information on mortgages taken out by households participating in the Survey on Household Income and Wealth (SHIW) carried out by the Bank of Italy every two years in the period 2006-2016. The dependent variable is the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is in table 9. Robust standard errors in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table 11: Variables from CR first information. Description and statistics

Variable	Description	N	Native Mean (S.D)	Immigrant Mean (S.D)	Diff.	t-stat
<i>Dependent variable</i>						
Approval $t-3d_t = 1$	Dummy variable, equal to 1 if bank i submitting a request for CR first information on natural person j grants a mortgage loan to j within the next three months, and 0 otherwise.	7,300,097	0.133 (0.34)	0.132 (0.34)	0.001	2.339**
Approval $t-6d_t = 1$	Dummy variable, equal to 1 if bank i submitting a request for CR first information on natural person j grants a mortgage loan to j within the next six months, and 0 otherwise.	7,300,097	0.1447 (0.35)	0.1423 (0.34)	0.0024	4.77***
<i>Mortgage Characteristics</i>						
Coint	Dummy variable, equal to 1 if there is co-borrower in loan contract, and 0 otherwise.	7,300,097	0.106 (0.31)	0.154 (0.36)	-0.047	-93.09***
<i>Borrower characteristics</i>						
Immigrant	Dummy variable, equal to 1 if the prospective mortgagor j about which bank i submit a request of CR first information is foreign-born, and 0 otherwise.	7,300,097	6,782,477	517,620		
Income*	The disposable income of the head of the household (in log).	7,300,097	22924.89 (6391)	15335 (4604)	7589.21	1.1e+03***
Age	Age of the prospective mortgagor j in years.	7,300,097	50.32 (12.48)	45.51 (10.07)	4.82	325.61***
Male	Dummy variable, equal to 1 if the prospective mortgagor j about which bank i submit a request of CR first information is male, and 0 otherwise.	7,300,097	0.645 (0.47)	0.644 (0.48)	0.001	2.21**
Credit history	Number of years since first time appearance of the prospective mortgagor in the Central Credit Register.	7,300,097	3.77 (5.7)	2.69 (4.6)	1.08	157.59***

Notes. Notes. Data source: Bank of Italy, Central Credit Register, 2015-2016. This table presents the description of variables, summary statistics, separately for immigrants and natives, and well as t-statistics for differences in mean.

Table 12: Loan approval regression results.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Immigrant	-0.025*** (0.000)	-0.027*** (0.001)	-0.024*** (0.000)	-0.026*** (0.001)	-0.025*** (0.000)	-0.027*** (0.001)	-0.026*** (0.000)	-0.028*** (0.001)
Age	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)
Age-squared	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Male	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)
Credit history	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Coint	0.325*** (0.001)	0.324*** (0.001)	0.325*** (0.001)	0.324*** (0.001)	0.325*** (0.001)	0.324*** (0.001)	0.325*** (0.001)	0.324*** (0.001)
ln (Income)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)
Immigrant × Credit history			-0.001*** (0.000)	-0.001*** (0.000)				
Immigrant × Main competitors					-0.005*** (0.002)	-0.005*** (0.002)		
Immigrant × Cooperative banks							0.014*** (0.002)	0.013*** (0.002)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216
Observations	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072

Notes. This table presents the coefficients from LPM regression for the request of CR first information made by bank in the period 2015-2016. The dependent variable is a dummy that takes the value 1 if a mortgage loans is granted conditional on banks' submitting a first information CR request within a 3-months period before the loan disbursement, and 0 otherwise. Description of explanatory variables is in table 11. ln(Income) is the logarithm of mortgagors' income. Income is an imputed variable obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Robust standard errors in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Appendix: additional tables

Table A1: Mortgage pricing regression results: baseline model with Bank $\times$ Time and Province $\times$ Time joint fixed effects, and clustered standard errors.

	Single-name mortgages			Joint mortgages		Full sample
	[1]	[2]	[3]	[4]	[5]	[6]
Immigrant	0.234*** (0.0342)	0.222*** (0.0323)				
ln (Loan amount)		-0.265*** (0.0170)	-0.265*** (0.0170)		-0.246*** (0.0200)	-0.257*** (0.0170)
Fixed		0.809*** (0.0712)	0.810*** (0.0712)		0.828*** (0.0786)	0.816*** (0.0747)
Age		-0.00800* (0.00442)	-0.00794* (0.00440)		-0.0202*** (0.00625)	-0.0141*** (0.00529)
Age-squared		9.02e-05* (4.65e-05)	8.96e-05* (4.64e-05)		0.000187*** (6.34e-05)	0.000137** (5.39e-05)
Male		0.0230*** (0.00720)	0.0227*** (0.00724)		0.0102*** (0.00321)	0.0128*** (0.00317)
Credit history		-0.0122*** (0.00261)	-0.0122*** (0.00261)		-0.00659*** (0.00204)	-0.00896*** (0.00206)
Europa			0.221*** (0.0338)			
Asia			0.296*** (0.0509)			
Africa			0.229*** (0.0319)			
North America			0.0663** (0.0301)			
Central America			0.253*** (0.0627)			
South America			0.203*** (0.0325)			
Oceania			0.0712 (0.0572)			
Cointimmi				0.406*** (0.0767)	0.379*** (0.0778)	0.457*** (0.0831)
Cointmix					0.113*** (0.0203)	0.197*** (0.0273)
Singleimmi						0.232*** (0.0348)
Cointnat						0.0890*** (0.0108)
Bank $\times$ Time FE	YES	YES	YES	YES	YES	YES
Province $\times$ Time FE	YES	YES	YES	YES	YES	YES
R-squared	0.318	0.403	0.403	0.348	0.438	0.418
Observations	490,189	490,189	490,189	464,152	464,152	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Credit Register in the period 2011-2016. The single-name mortgage sample includes loans in name of a single borrower; the joint mortgage sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. Robust standard errors clustered standard errors at the bank level are in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table A2: Mortgage pricing regression results: baseline model for Loan amount $\geq$ €75,000.

	Single-name mortgages			Joint mortgages		Full sample
	[1]	[2]	[3]	[4]	[5]	[6]
Immigrant	0.247*** (0.00608)	0.236*** (0.00563)				
Immigrant		0.227*** (0.00572)				
ln (Loan amount)		-0.340*** (0.00358)	-0.340*** (0.00358)		-0.311*** (0.00395)	-0.327*** (0.00265)
Fixed		0.749*** (0.00368)	0.749*** (0.00368)		0.765*** (0.00352)	0.754*** (0.00255)
Age		-0.00708*** (0.000963)	-0.00701*** (0.000963)		-0.0191*** (0.000906)	-0.0131*** (0.000659)
Age-squared		8.04e-05*** (1.13e-05)	7.98e-05*** (1.13e-05)		0.000172*** (1.04e-05)	0.000125*** (7.65e-06)
Male		0.0236*** (0.00277)	0.0233*** (0.00277)		0.0105*** (0.00264)	0.0128*** (0.00190)
Credit history		-0.0111*** (0.000358)	-0.0111*** (0.000358)		-0.00594*** (0.000381)	-0.00791*** (0.000258)
Europa			0.227*** (0.00696)			
Asia			0.300*** (0.0168)			
Africa			0.230*** (0.0195)			
North America			0.0762** (0.0317)			
Central America			0.259*** (0.0483)			
South America			0.205*** (0.0169)			
Oceania			0.0512 (0.0588)			
Cointimmi				0.427*** (0.00538)	0.390*** (0.00502)	0.482*** (0.00492)
Cointmix					0.115*** (0.00499)	0.213*** (0.00499)
Singleimmi						0.237*** (0.00567)
Cointnat						0.102*** (0.00205)
Time FE	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
R-squared	0.288	0.374	0.375	0.315	0.405	0.389
Observations	466,785	466,785	466,785	449,503	449,503	916,288

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Credit Register in the period 2011-2016. The single-name mortgage sample includes loans in name of a single borrower; the joint mortgage sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. Robust standard errors are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table A3: Mortgage pricing regression results: augmented model with imputed LTV ratio and mortgagors' income, and with Bank×Time joint fixed effects, Province×Time joint fixed effects and clustered standard errors.

	Single-name mortgages			Joint mortgages		Full sample
	[1]	[2]	[3]	[4]	[5]	[6]
Immigrant	0.193*** (0.0294)	0.172*** (0.0281)				
ln (Loan amount)	-0.291*** (0.0311)	-0.258*** (0.0323)	-0.257*** (0.0322)	-0.343*** (0.0286)	-0.308*** (0.0293)	-0.277*** (0.0291)
Fixed	0.829*** (0.0697)	0.820*** (0.0699)	0.820*** (0.0699)	0.844*** (0.0773)	0.818*** (0.0774)	0.818*** (0.0734)
Age	-0.00849* (0.00444)	0.00843** (0.00413)	0.00849** (0.00411)	-0.0236*** (0.00673)	-0.00242 (0.00737)	0.00142 (0.00563)
Age-squared	0.000101** (4.69e-05)	-6.00e-05 (4.46e-05)	-6.06e-05 (4.45e-05)	0.000244*** (7.02e-05)	4.40e-05 (7.66e-05)	2.54e-06 (5.90e-05)
Male	0.0247*** (0.00806)	0.0983*** (0.00943)	0.0979*** (0.00944)	0.00990*** (0.00314)	0.0148*** (0.00299)	0.0542*** (0.00424)
Credit history	-0.0121*** (0.00252)	-0.0119*** (0.00250)	-0.0119*** (0.00250)	-0.00574*** (0.00193)	-0.00534*** (0.00193)	-0.00804*** (0.00195)
LTV	0.00232 (0.00197)	0.00192 (0.00195)	0.00189 (0.00194)	0.00854*** (0.00275)	0.00784*** (0.00274)	0.00413* (0.00221)
ln(Income)		-0.333*** (0.0317)	-0.332*** (0.0317)		-0.414*** (0.0300)	-0.341*** (0.0243)
Europa			0.171*** (0.0294)			
Asia			0.248*** (0.0466)			
Africa			0.178*** (0.0289)			
North America			0.0145 (0.0344)			
Central America			0.202*** (0.0584)			
South America			0.150*** (0.0265)			
Oceania			0.0349 (0.0620)			
Cointimmi				0.274*** (0.0591)	0.243*** (0.0579)	0.331*** (0.0630)
Cointmix				0.113*** (0.0197)	0.110*** (0.0195)	0.151*** (0.0256)
Singleimmi						0.149*** (0.0324)
Cointnat						0.0435*** (0.0101)
Time × bank FE	YES	YES	YES	YES	YES	YES
Time × Province FE	YES	YES	YES	YES	YES	YES
R-squared	0.414	0.416	0.416	0.450	0.454	0.432
Observations	490,189	490,189	490,189	464,152	464,152	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy' Credit Register in the period 2011-2016. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is in table 2. LTV is the ratio between the amount of the mortgage and the lender-assessed value of the house. Income is the mortgagors' income. Both LTV and Income are imputed variables obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Robust standard errors clustered standard errors at the bank level are in parentheses. ***p < 0.001, **p < 0.01, *p < 0.05.

Table A4: Interest rate regression underlying the Oaxaca-Blinder decomposition of Table 6

	Single name mortgages		Joint mortgages		Full sample	
	[1]	[2]	[3]	[4]	[5]	[6]
ln (Loan amount)	0.0118*** (0.000688)	-0.122 (0.205)	0.0284*** (0.000701)	0.947*** (0.241)	0.0184*** (0.000423)	0.252 (0.156)
Fixed	-0.00720*** (0.000881)	0.00888** (0.00426)	-0.0129*** (0.000774)	-0.00806** (0.00370)	-0.00984*** (0.000582)	-0.000842 (0.00279)
Age	0.00385*** (0.00114)	0.658*** (0.230)	0.107*** (0.00350)	1.354*** (0.201)	0.0470*** (0.00172)	0.971*** (0.152)
Age-squared	-0.00632*** (0.00114)	-0.322*** (0.117)	-0.0999*** (0.00338)	-0.577*** (0.0987)	-0.0468*** (0.00167)	-0.454*** (0.0758)
Male	-0.00173*** (0.000277)	-0.00109 (0.00761)	2.62e-05 (5.74e-05)	-0.00433 (0.00644)	-0.000220** (0.000108)	0.00218 (0.00489)
Credit history	0.0265*** (0.00100)	0.00223 (0.00422)	0.0258*** (0.000687)	-0.000547 (0.00136)	0.0358*** (0.000544)	-0.0110*** (0.00193)
Observations	490,189	490,189	432,191	432,191	922,380	922,380

Notes. This table shows the OLS regression results generated during the Oaxaca blinder decomposition analysis using the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016, which decomposes the difference in the mean of mortgage rates between immigrants and natives into two components: the explained component due to differences in the coefficients, and the unexplained component due to differences in the characteristics of the two groups. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. Robust standard errors are in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table A5: Oaxaca-Blinder decomposition of mortgage interest rate gap between immigrants and Italian natives with clustered standard errors at bank level

	Single name mortgages		Joint mortgage		Full sample	
	Mean interest rate	Decomposition	Mean interest rate	Decomposition	Mean interest rate	Decomposition
Immigrants	3.340*** (0.0358)		3.579*** (0.0817)		3.474*** (0.0582)	
Natives	3.157*** (0.0411)		3.261*** (0.0414)		3.205*** (0.0408)	
Differential	0.184*** (0.0295)		0.319*** (0.0866)		0.269*** (0.0623)	
Explained components		0.0269*** (0.00592)		0.0479*** (0.0108)		0.0443*** (0.00849)
Unexplained components		0.157*** (0.0282)		0.271*** (0.0830)		0.225*** (0.0594)
Observations	490,189	490,189	432,191	432,191	922,380	922,380

Notes. This table shows results from Oaxaca-Blinder decomposition of the mean interest rate differential between mortgages made out to immigrants and native-born Italians. The explained component (or endowments effect) is the part of the interest rate differential that is attributable to group differences in explanatory variables. The unexplained component (or coefficients effect) is the part of the interest rate differential that is attributable to differences in coefficients of explanatory variables. Robust standard errors clustered standard errors at the bank level are in parentheses. *** p -value < 0.01, ** p -value < 0.05, * p -value < 0.1.

Table A6: Mortgage pricing regression results: borrowers' credit history, dominant lenders and cooperative banks (full sample)

	Full Sample			
	[1]	[2]	[3]	[4]
Singleimmi	0.2628*** (0.0062)	0.2642*** (0.0062)	0.2485*** (0.0066)	0.2592*** (0.0061)
Cointimmi	0.4895*** (0.0050)	0.4878*** (0.0050)	0.4040*** (0.0058)	0.5192*** (0.0054)
Cointmix	0.2077*** (0.0050)	0.2060*** (0.0050)	0.2077*** (0.0050)	0.2075*** (0.0050)
Cointnat	0.0951*** (0.0020)	0.0936*** (0.0020)	0.0953*** (0.0020)	0.0949*** (0.0020)
Credit history	-0.0095*** (0.0003)			
Singleimmi $\times$ Credit history	-0.0082*** (0.0016)			
Cointimmi $\times$ Credit history	-0.0002 (0.0025)			
$\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.0920*** (0.0024)		
Singleimmi $\times$ $\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.0745*** (0.0136)		
Cointimmi $\times$ $\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		0.0010 (0.0224)		
Singleimmi $\times$ Main competitors			-0.0099 (0.0120)	
Cointimmi $\times$ Main competitors			0.2306*** (0.0100)	
Singleimmi $\times$ Cooperative				-0.0909*** (0.0145)
Cointimmi $\times$ Cooperative				-0.2165*** (0.0119)
Time FE	YES	YES	YES	YES
Prov FE	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES
R-squared	0.3827	0.3827	0.3829	0.3828
Observations	954,341	954,341	954,341	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of moderators is reported in table 2. All columns include the same control variables of specifications [6] in Table 4. Robust standard errors are in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table A7: Mortgage pricing regression results: borrowers' credit history, dominant lenders and cooperative banks with Bank×Time and Province×Time joint fixed effects, and clustered standard errors.]

	Single-name mortgages				Joint mortgages				Full sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Immigrant	0.2322*** (0.0370)	0.2338*** (0.0375)	0.2152*** (0.0245)	0.2351*** (0.0361)					0.2450*** (0.0404)	0.2465*** (0.0410)	0.2287*** (0.0240)	0.2434*** (0.0381)
Singleimmi									0.1969*** (0.0273)	0.1955*** (0.0273)	0.1970*** (0.0274)	0.1967*** (0.0273)
Cointmix									0.0891*** (0.0109)	0.0878*** (0.0109)	0.0894*** (0.0110)	0.0889*** (0.0108)
Cointnat									0.4569*** (0.0842)	0.4558*** (0.0841)	0.3617*** (0.0288)	0.4834*** (0.0893)
Cointimmi					0.3792*** (0.0790)	0.3794*** (0.0791)	0.2905*** (0.0240)	0.4039*** (0.0838)				
Credit history	-0.0120*** (0.0026)				-0.0066*** (0.0020)				-0.0088*** (0.0021)			
Immigrant x Credit history	-0.0047 (0.0035)											
Cointimmi x Credit history					-0.0006 (0.0040)				0.0004 (0.0038)			
Singleimmi x Credit history									-0.0061 (0.0038)			
Credit history _t 3years		-0.1123*** (0.0225)				-0.0635*** (0.0178)				-0.0847*** (0.0175)		
Immigrant x Credit history _t 3years		-0.0431 (0.0331)										
Cointimmi x Credit history _t 3yrs						-0.0093 (0.0350)				0.0019 (0.0319)		
Singleimmi x Credit history _t 3yrs										-0.0569 (0.0373)		
Immigrant x Main competitors			0.0223 (0.0904)									
Cointimmi x Main competitors							0.2352* (0.1392)				0.2570* (0.1518)	
Singleimmi x Main competitors											0.0085 (0.0805)	
Immigrant x Cooperative				-0.0866* (0.0461)								-0.1909** (0.0906)
Cointimmi x Cooperative								-0.1755** (0.0850)				-0.0756* (0.0421)
Singleimmi x Cooperative												
Time X bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time X Province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.4032	0.4032	0.4032	0.4032	0.4377	0.4376	0.4383	0.4379	0.4179	0.4178	0.4182	0.4180
Observations	490,189	490,189	490,189	490,189	464,152	464,152	464,152	464,152	954,341	954,341	954,341	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLIR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of moderators is reported in table 2. All columns include the same control variables of specifications [2], [5], and [6] in Table 4. Robust standard errors clustered standard errors at the bank level are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table A8: Mortgage pricing regression results- Augmented model: borrowers' credit history, dominant lenders and cooperative banks

	Single-name mortgages				Joint mortgages				Full sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Immigrant	0.2149*** (0.0072)	0.2144*** (0.0072)	0.1980*** (0.0075)	0.2145*** (0.0070)								
Singleimm												
Cointimm					0.2839*** (0.0062)	0.2836*** (0.0062)	0.2060*** (0.0069)	0.3126*** (0.0065)	0.1926*** (0.0065)	0.1935*** (0.0065)	0.1780*** (0.0069)	0.1894*** (0.0064)
Cointmix									0.3846*** (0.0056)	0.3824*** (0.0056)	0.3003*** (0.0063)	0.4146*** (0.0059)
Cointnat									0.1674*** (0.0048)	0.1662*** (0.0048)	0.1674*** (0.0048)	0.1671*** (0.0048)
Credit history	-0.0125*** (0.0004)				-0.0058*** (0.0004)				0.0545*** (0.0021)	0.0535*** (0.0021)	0.0547*** (0.0021)	0.0543*** (0.0021)
Immigrant × Credit history	-0.0069*** (0.0015)								-0.0084*** (0.0003)			
$\mathbb{I}(\text{Credit history} > 3 \text{ yrs})$		-0.1147*** (0.0031)				-0.0547*** (0.0037)				-0.0794*** (0.0024)		
Immigrant × $\mathbb{I}(\text{Credit history} > 3 \text{ yrs})$		-0.0647*** (0.0132)										
Immigrant × Main competitors			0.0055 (0.0117)									
Immigrant × Cooperative				-0.1021*** (0.0142)								
Cointimm × Credit history					-0.0006 (0.0023)				0.0001 (0.0023)			
Cointimm × $\mathbb{I}(\text{Credit history} > 3 \text{ yrs})$						-0.0086 (0.0214)				-0.0002 (0.0214)		
Cointimm × Main competitors							0.2073*** (0.0096)				0.2265*** (0.0094)	
Cointimm × Cooperative								-0.2026*** (0.0117)				-0.2208*** (0.0114)
Singleimm × Credit history									-0.0078*** (0.0015)			
Singleimm × $\mathbb{I}(\text{Credit history} > 3 \text{ yrs})$										-0.0746*** (0.0131)		
Singleimm × Main competitors											-0.0082 (0.0114)	
Singleimm × Cooperative												-0.0914*** (0.0138)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Prov. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.4071	0.4070	0.4070	0.4071	0.4451	0.4450	0.4455	0.4453	0.4243	0.4242	0.4245	0.4244
Observations	490,189	490,189	490,189	490,189	464,152	464,152	464,152	464,152	954,341	954,341	954,341	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASLR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. LTV is the ratio between the amount of the mortgage and the lender self-assessed value of the house. Income is the mortgages' income. Both LTV and Income are imputed variables obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". All columns include the same control variables of specifications [2], [5] and [6] in Table 5. Robust standard errors are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table A9: Mortgage pricing regression results- Augmented model: borrowers' credit history, dominant lenders and cooperative banks with Bank×Time and Province×Time joint fixed effects, and clustered standard errors.

	Single-name mortgages				Joint mortgages				Full sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Immigrant	0.1803*** (0.0325)	0.1801*** (0.0332)	0.1644*** (0.0337)	0.1849*** (0.0319)	0.2424*** (0.0592)	0.2424*** (0.0591)	0.1552*** (0.0418)	0.2678*** (0.0642)	0.3314*** (0.0644)	0.3294*** (0.0640)	0.2372*** (0.0401)	0.3583*** (0.0701)
Cointimmi									0.1587*** (0.0375)	0.1599*** (0.0383)	0.1444*** (0.0339)	0.1601*** (0.0360)
Sing'eimmi									0.1507*** (0.0256)	0.1494*** (0.0258)	0.1508*** (0.0258)	0.1505*** (0.0256)
Cointmix									0.0437*** (0.0101)	0.0425*** (0.0102)	0.0440*** (0.0103)	0.0435*** (0.0101)
Cointnat									-0.0080*** (0.0020)			
Credit history	-0.0118*** (0.0025)				-0.0053*** (0.0019)							
Immigrant × Credit history	-0.0037 (0.0034)											
Cointimmi × Credit history					0.0004 (0.0040)				0.0008 (0.0038)			
Sing'eimmi × Credit history									-0.0047 (0.0038)			
$\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.1103*** (0.0216)				-0.0505*** (0.0168)				-0.0765*** (0.0164)		
Immigrant × $\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$		-0.0367 (0.0329)										
Cointimmi × $\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$						-0.0047 (0.0342)				0.0032 (0.0317)		
Sing'eimmi × $\mathbb{1}(\text{Credit history} > 3 \text{ yrs})$										-0.0481 (0.0374)		
Immigrant × Main competitors			0.0241 (0.0906)									
Cointimmi × Main competitors							0.2317* (0.1374)				0.2544* (0.1517)	
Sing'eimmi × Main competitors											0.0109 (0.0803)	
Immigrant × Cooperative				-0.0877* (0.0473)								-0.0768* (0.0431)
Sing'eimmi × Cooperative								-0.1760** (0.0830)				-0.1931** (0.0896)
Cointimmi × Cooperative									YES (0.0000)	YES (0.0000)	YES (0.0000)	YES (0.0000)
Time X bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time X Province FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.4159	0.4160	0.4159	0.4160	0.4539	0.4539	0.4545	0.4541	0.4316	0.4316	0.4320	0.4318
Observations	490,189	490,189	490,189	490,189	464,152	464,152	464,152	464,152	954,341	954,341	954,341	954,341

Notes. This table presents the coefficients from OLS regressions for mortgages to natural persons included in the Bank of Italy's Central Credit Register and ASIR survey in the period 2011-2016. The single-name mortgages sample includes loans in name of a single borrower; the joint mortgages sample includes loans made out to two co-borrowers; the full sample includes both single-name and joint mortgages. The dependent variable is *APR*, the annual interest rate charged to borrowers on granted mortgages in percentage points. Description of explanatory variables is reported in table 2. LTV is the ratio between the amount of the mortgage and the lender self-assessed value of the house. Income is the mortgagors' income. Both LTV and Income are imputed variables obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". All columns include the same control variables of specifications [2], [5] and [6] in Table 5. Robust standard errors clustered standard errors at the bank level are in parentheses. ****p*-value < 0.001, ***p*-value < 0.01, **p*-value < 0.05.

Table A10: Loan approval regression results. (Dep. var.: Approval | $t_{-6}d_t = 1$).

	[1])	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Immigrant	-0.027*** (0.000)	-0.029*** (0.001)	-0.025*** (0.000)	-0.027*** (0.001)	-0.026*** (0.000)	-0.028*** (0.001)	-0.028*** (0.000)	-0.030*** (0.001)
Age	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)
Age-squared	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Male	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)
Credit history	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
Coint	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)	0.342*** (0.001)
ln(Income)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)
Immigrant×Credit history			-0.000*** (0.000)	-0.000*** (0.000)				
Immigrant×Main competitors					-0.003 (0.002)	-0.003 (0.002)		
Immigrant×Cooperative banks							0.013*** (0.002)	0.013*** (0.002)
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.226	0.226	0.226	0.226	0.226	0.226	0.226	0.226
Observations	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072	7,300,072

Notes. This table presents the coefficients from LPM regression for the request of CR first information made by bank in the period 2015-2016. The dependent variable is a dummy that takes the value 1 if a mortgage loans is granted conditional on banks' submitting a first information CR request within a 6-months period before the loan disbursement, and 0 otherwise. Description of explanatory variables is in table 11. ln(Income) is the logarithm of mortgagors' income. Income is an imputed variable obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Robust standard errors in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table A11: Loan approval regression results. (Dep. var.: Approval | $t_{-3}d_t = 1$).

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Immigrant	-0.025*** (0.000)	-0.027*** (0.001)	-0.023*** (0.000)	-0.026*** (0.001)	-0.024*** (0.000)	-0.027*** (0.001)	-0.026*** (0.000)	-0.028*** (0.001)
Age	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)
Age-squared	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Male	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)	0.016*** (0.000)	0.018*** (0.000)
Credit history	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Coint	0.324*** (0.001)	0.323*** (0.001)	0.324*** (0.001)	0.323*** (0.001)	0.324*** (0.001)	0.323*** (0.001)	0.324*** (0.001)	0.323*** (0.001)
ln(Income)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)
Immigrant×Credit history			-0.001*** (0.000)	-0.001*** (0.000)				
Immigrant×Main competitors					-0.005*** (0.002)	-0.005*** (0.002)		
Immigrant×Cooperative banks							0.013*** (0.002)	0.013*** (0.002)
Bank×Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Province×Time FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.220	0.220	0.220	0.220	0.220	0.220	0.220	0.220
Observations	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490

Notes. This table presents the coefficients from LPM regression for the request of CR first information made by bank in the period 2015-2016. The dependent variable is a dummy that takes the value 1 if a mortgage loans is granted conditional on banks' submitting a first information CR request within a 3-months period before the loan disbursement, and 0 otherwise. Description of explanatory variables is in table 11. ln(Income) is the logarithm of mortgagors' income. Income is an imputed variable obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Regressions include bank×time and province×time fixed effects. Robust standard errors in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

Table A12: Loan approval regression results. (Dep. var.: Approval | $t_{-6}d_t = 1$).

	[1])	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Immigrant	-0.026*** (0.000)	-0.029*** (0.001)	-0.025*** (0.001)	-0.027*** (0.001)	-0.026*** (0.000)	-0.028*** (0.001)	-0.027*** (0.000)	-0.029*** (0.001)
Age	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)	-0.011*** (0.000)
Age-squared	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
Male	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)	0.017*** (0.000)	0.019*** (0.000)
Credit history	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.000)
Coint	0.342*** (0.001)	0.341*** (0.001)	0.342*** (0.001)	0.341*** (0.001)	0.342*** (0.001)	0.341*** (0.001)	0.342*** (0.001)	0.341*** (0.001)
ln(Income)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)		-0.005*** (0.001)
Immigrant×Credit history			-0.000*** (0.000)	-0.000*** (0.000)				
Immigrant×Main competitors					-0.003 (0.002)	-0.003 (0.002)		
Immigrant×Cooperative banks							0.013*** (0.002)	0.013*** (0.002)
Bank×Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Province×Time FE	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.229	0.229	0.229	0.229	0.229	0.229	0.229	0.229
Observations	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490	7,299,490

Notes. This table presents the coefficients from LPM regression for the request of CR first information made by bank in the period 2015-2016. The dependent variable is a dummy that takes the value 1 if a mortgage loans is granted conditional on banks' submitting a first information CR request within a 6-months period before the loan disbursement, and 0 otherwise. Description of explanatory variables is in table 11. ln(Income) is the logarithm of mortgagors' income. Income is an imputed variable obtained by using the Random Forest algorithm trained on the Banks of Italy's SHIW dataset with the R package "Random Forest". Regressions include bank×time and province×time fixed effects. Robust standard errors in parentheses. *** p -value < 0.001, ** p -value < 0.01, * p -value < 0.05.

The geopolitics of Aid: U.S. and China’s strategic financial flows

Emmanuel Caiazzo*

Pietro Panizza[†]

Alberto Zazzaro[‡]

Abstract

The competition between the United States (U.S.) and China is poised to influence the course of history in the near future. Is international aid another ground of confrontation? By analyzing the Creditor Reporting System dataset by OECD and the Global Chinese Development Finance Dataset over a 20-year period (2000-2019), we investigate whether aid commitment allocation strategies reflect competitive dynamics or attempts at sphere-of-influence consolidation. Employing an instrumental variable approach, we document that the U.S. increases its aid commitment in recipient countries where China commits more funds. Moreover, we document that the American response to Chinese financing is greater in the presence of political or commercial ties to the U.S. and the recipient country, and increased following the Xi Jinping’s election. Overall, our findings are consistent with the view that U.S. international aid allocation criteria were oriented to counter the expansion of the Chinese influence.

Keywords— International aid, Donor coordination, US-Sino competition

JEL Code— F35, O19

*University of Naples Parthenope. emmanuel.caiazzo@unina.it.

[†]University of Naples Federico II, Italy

[‡]University of Naples Federico II, CSEF and MoFiR. alberto.zazzaro@unina.it.

1 Introduction

In recent decades, the U.S. has gradually shifted its international relationship with China. In 1998, during a visit to Peking University, President Clinton referred to the 21st century as China’s century.¹ Just over twenty years later, the first Trump administration defined China as a country hostile to U.S. interests and launched a fierce trade war with Beijing (McFaul, 2020; Fajgelbaum et al., 2024). President Biden, in his inauguration speech, emphasized that the U.S. and China are “*in competition for the 21st century*”.² The Sino–US rivalry now extends to many fields, such as innovation and technology (Kennedy and Lim, 2018), trade (Hopewell, 2020), and military (Allison, 2017). The scope of this paper is to examine whether and to what extent China and the United States compete in the realm of foreign aid.

A large body of literature indicates that foreign aid has been used as a powerful policy instrument through which donor countries aim to enhance their soft power vis-à-vis recipient countries (Morgenthau, 1962; Thacker, 1999; Calomiris, 2000; Whang et al., 2019). In this perspective, since the late 1990s, China has progressively expanded its foreign aid disbursements. The Belt and Road Initiative (BRI), launched after Xi Jinping’s ascent to power, constitutes a massive, multi-year investment campaign with the declared aim of boosting China’s political influence worldwide.³ To give a sense of the scale of the BRI, in inflation-adjusted terms the Marshall Plan amounted to roughly 10 percent of the total value of BRI financing. (Shambaugh, 2015; Natsios, 2020).

Does U.S. foreign aid allocation respond to the intensification of China’s foreign aid programs, and if so how? Political scientists and observers of international affairs broadly agree that foreign aid constitutes a central arena of geopolitical competition between China and the United States (Edel and Brands, 2019; Tierney, 2024).

However, from a theoretical perspective, it is not clear if American and Chinese aid are strategic substitutes or strategic complements. To the extent that international aid programs provide private benefits to donors (Dudley and Montmarquette, 1976), by establishing silent or explicit agreements to define their spheres of influence, the U.S. and China can effectively coordinate their actions and increase rents from aid, mirroring collusion equilibria of the oligopoly theory (Steinwand, 2015). In this case, Chinese and U.S. aid are strategic substitutes. However, the desire to limit the expansion of the rival’s influence can lead to coordination failures. In this case, Chinese and American aid are strategic complements.

Considering the substantial geopolitical implications of international aid and the detrimental effects of aid fragmentation among donors on aid’s effectiveness (Easterly, 2007; Knack and Rahman, 2007; Djankov et al., 2009; Fuchs et al., 2015), exploring whether Chinese aid and

¹“*In the 21st century – your century – China and the United States will face the challenge of security in Asia. On the Korean Peninsula, where once we were adversaries, today we are working together for a permanent peace and a future freer of nuclear weapons*”. Full speech is available at the [White House Archive](#).

²See also Josh Rogin on [Washington Post](#), 10 Feb. 2022.

³Xi Jinping stated in 2014, “*We should increase China’s soft power, craft a compelling Chinese narrative, and better communicate China’s message to the world*”.

American aid function as strategic substitutes or strategic complements is a significant aspect that our paper seeks to address.

We collect data on Chinese and U.S. international aid from [Goodman et al. \(2024\)](#) and the OECD website, respectively. The sample period covers 20 years (2000-2019), during which China and the United States initiated a total of approximately 45,000 projects worldwide, involving a total of about 900 billion dollars.⁴

From an empirical standpoint, the main issue is the presence of reverse causality. To address this problem, we adopt an IV strategy based on the work of [Dreher et al. \(2021\)](#). That is, we instrument the amount of dollars China commits to spend in a recipient country in a given year by using a composite term that reflects the annual Chinese domestic overproduction of industrial project inputs—cement, steel, glass, iron, and aluminum—and the likelihood that a recipient country will have a diplomatic relationship with Taiwan over the sample period. The underlying premise behind the first component of the instrument is that China’s annual production of raw materials plays a crucial role in determining the availability of financing from the Chinese government for foreign infrastructure projects. The Chinese government regards domestic production inputs, such as steel and aluminum, as strategically important commodities and consequently produces these materials in quantities that significantly exceed domestic demand. As a result, in years with high production volumes of project inputs, an increase in the availability of Chinese government financing for economic infrastructure initiatives abroad is expected ([Dreher et al., 2021](#); [Bluhm et al., 2025](#)).

The second component of the instrument relies on an extensive literature identifying political alliances between the donor and the recipient as the main determinant of international aid. In this perspective, the fact that a recipient country does not have a diplomatic relationship with Taiwan can be interpreted as the recipient’s support for the *one China policy*. This policy emphasizes the recognition of the People’s Republic of China as the sole legitimate government of China and typically entails the severance of formal diplomatic ties with Taiwan. Consequently, the absence of such ties serves as a proxy indicator of the recipient’s political alignment with Beijing, reflecting broader strategic and geopolitical considerations that influence aid allocation decisions.

Our IV regression reveals a positive association between Chinese and American aid financing. That is, in countries where China commits more funds, the U.S. also commits more resources. To the extent that international aid is a tool used by donor countries to gain political influence over the recipient country ([Wellner et al., 2025](#)), we interpret our findings within the framework of competition between China and the United States in establishing their own spheres of influence around the world. In fact, a plausible rationale for the result we document is that the U.S. is attempting to counter the rise of China’s soft power. To support this interpretation, we

⁴In the paper, we will consider the amount of dollars that countries commit to disbursing rather than the amount of dollars actually spent. Focusing on commitments rather than expenditures is a widely used approach in the literature, since commitment data are less influenced by demand-side fluctuations.

show that the U.S. response—by increasing aid—is stronger when Chinese aid is distributed in countries with strong political or commercial ties to the United States. Furthermore, we observe that the U.S. response to Chinese aid increased after Xi Jinping’s election, starting when the expansion of soft power became a priority for the Chinese government explicitly (Li, 2022).

2 Related literature

The paper relates to an extensive literature discussing determinants of international aid. The stated intention of international aid is to accelerate economic growth in developing nations (Rosenstein-Rodan, 1961). Nonetheless, debates persist regarding their actual effectiveness in achieving this purpose (Bauer, 1976; Burnside and Dollar, 2000; Easterly, 2003; Easterly et al., 2004; Easterly, 2009). Countries continue to dedicate significant resources to international aid programs annually. In 2023, members of the Development Assistance Committee (DAC) contributed a total of USD 223.7 billion in official development assistance (ODA), representing 0.37% of their combined gross national income. This marked the fifth year in a row that ODA reached an all-time high.

Political and strategic interests often guide foreign aid policies. As early as the 1960s, Hans Morgenthau 1962 emphasized the importance of the United States establishing a foreign aid policy.⁵

Donors frequently utilize aid programs to secure political concessions from recipient nations or to reinforce their international political alliances (Morgenthau et al., 1985; Thacker, 1999; Calomiris, 2000; Lebovic, 2005; Whang et al., 2019). Evidence from Alesina and Dollar (2000) suggests that previous colonial relationships between donors and recipients play a crucial role in the determination of aid provision. Additionally, research shows that higher exports from donor countries to recipient nations are associated with increased aid allocations (Dudley and Montmarquette, 1976; Berthélemy et al., 2006; Dreher et al., 2015). Furthermore, Younas (2008) suggests that donor countries are more likely to allocate aid to recipient countries that import goods in which donors have a comparative advantage in production, particularly capital goods.⁶

A substantial body of research links international aid to the quality of institutions in recipient countries. The reasoning behind this connection is that weak economic and political institutions increase the risk of aid being captured by local elites, thereby limiting its effectiveness

⁵Ibid: “*It is in fact pointless even to raise the question whether the United States ought to have a policy of foreign aid — just as it is to ask whether the United States should have a foreign political or military policy. The U.S. has interests abroad that cannot be secured solely through military means, and traditional diplomatic methods are only partially suitable for supporting those interests. Without foreign aid, these interests would lack support at all.*”

⁶Evidence suggests that the political interests of key shareholders in institutions like the IMF and World Bank influence how aid funds are allocated. See, among others, Lang and Presbitero (2018) and Aiyar et al. (2024).

in promoting economic growth (Boone, 1996). However, when evaluating whether less corrupt governments receive more bilateral aid, Alesina and Weder (2002) conclude that corrupt governments often receive similar amounts of aid as less corrupt ones. Similarly, Svensson (1999) provide evidence indicating that aid is not necessarily directed toward more democratic nations. On the issue of human rights abuses, anecdotal evidence suggests that some donors suspend aid when recipient countries commit severe human rights violations (O’Keefe, 2007).

Other studies explore whether international aid can enhance the quality of institutions in recipient countries, based on the idea that donors can influence recipient governments. In this perspective, some papers demonstrate that recipient countries modify their policies by adopting democratic reforms in response to international aid programs (Bermeo, 2011; Kersting and Kilby, 2014). In contrast, Knack (2004) finds no indication that aid consistently fosters long-term democratic development. This mixed evidence has theoretical roots: on one hand, by supporting economic and social development, aid may lay the groundwork for democratization (Lipset, 1959); on the other hand, aid can provide recipient governments with greater access to material resources and often strengthen regime legitimacy (Dasandi and Erez, 2019).

Theoretical research has examined the criteria guiding international aid allocation and the strategic interactions among donors, primarily to better understand the causes of aid fragmentation. Steinwand (2015) introduce a conceptual framework that distinguishes between cooperation and competition among donors, considering factors such as lead donorship and whether aid functions as a private or public good. Annen and Moers (2017) demonstrate that aid fragmentation can emerge as an inefficient equilibrium when donors aim to maximize the relative impact of their individual programs.

To our knowledge, the competition between the United States and China in the realm of international aid has received limited attention in the existing literature. Chung et al. (2023) find that U.S. citizens who are aware of America’s competition with China in aid efforts are more willing to support increased government spending. Conversely, Blair et al. (2022) report that Chinese aid to Africa does not boost—and may even decrease—local support for China. In contrast, U.S. aid appears to bolster support for the United States and reinforce recipients’ commitment to liberal democratic values, such as the importance of elections. Chinese aid, on the other hand, does not seem to undermine this commitment and may even strengthen it.

3 Empirical analysis

3.1 International aid

Data on Chinese international aid are sourced from Goodman et al. (2024). The dataset includes 20,985 projects in 165 low- and middle-income countries, funded by loans and grants from Chinese official sector institutions. It spans 22 commitment years (2000-2021) and encompasses various regions, sectors, and types of financial and in-kind transfers. This aligns with the OECD

definition of official financial flows, including Official Development Assistance (ODA) and Other Official Flows (OOF).

Projects are categorized as ODA if they meet three criteria:

- (i) The primary goal is to promote economic development and welfare in the recipient country.
- (ii) The project occurs in a country eligible for ODA based on its income level.
- (iii) The supporting official commitment is concessional, such as grants, technical assistance, or loans with a grant element.

Projects that receive official support but do not meet all three criteria are classified as OOF. If insufficient information prevents reliable categorization as ODA or OOF, projects are labeled as Vague.

Each project/activity is classified into one of five categories: Development, Commercial, Representational, Mixed, or Military.

- (i) Development projects focus on promoting economic development and welfare in the recipient country.
- (ii) Commercial projects aim to advance the commercial interests of the originating country, such as promoting exports of Chinese goods and services.
- (iii) Mixed projects have multiple purposes and cannot be solely identified as development, commercial, or representational.
- (iv) Military projects aim to advance the security interests of the originating country or strengthen military capabilities in the recipient country.

From the original dataset, to better compare with the data for US aid discussed below, we retain only projects with a development intention, which represent approximately 77% of the total. Moreover, we exclude Covid-19 years from the sample. This leaves us with around 16,000 projects, of which almost 90% are categorized as ODA. In our baseline, we consider the totality of the aid. In a robustness check, we focus on the aid classified as ODA.

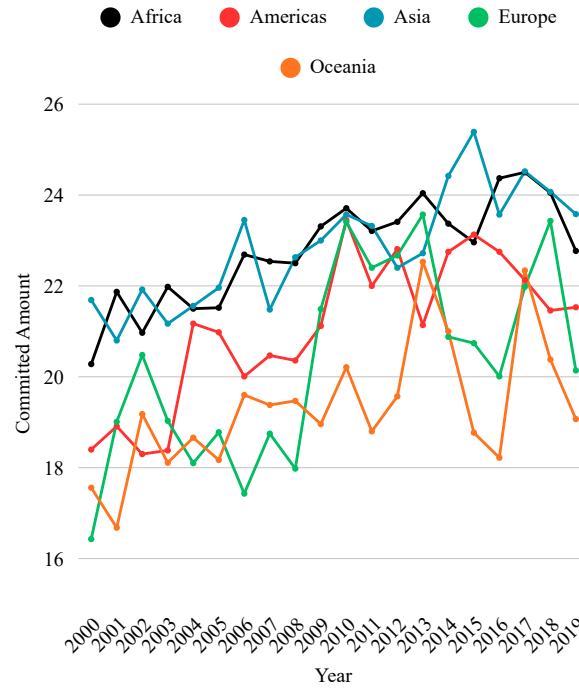
Data on US international aid are sourced from the Creditor Reporting System of the OECD. We consider projects/activities over the 20 years 2000-2019. The dataset tracks around 26,000 projects, more than 99% categorized as ODA.

Aid activities of both China and the U.S. can be summarized through a series of graphs and maps. Figure 1 provides a broad overview of Chinese international aid. Panel 1a shows the amount of USD committed by China (on a log scale) during the sample period across different regions of the world. Africa and Asia are the regions where most of the funds are allocated. Panel 1b displays China's total financing over the years. On average, the amount of USD committed by China in international aid grows at an average rate of 11% per year. Aid activities for the U.S. present a similar picture (Figure 2). Also for the U.S., Asia and Africa are the regions where most of the funds are allocated. However, over the sample period, the amount of USD committed by the U.S. exhibits a much more modest growth rate compared to China (5% yearly average).

We can graphically represent the similarities between China's and the U.S.'s aid allocation

Figure 1: Chinese International Aid

(a) Committed amount of dollars (log) per year and by region



(b) Committed amount of dollars (log) per year

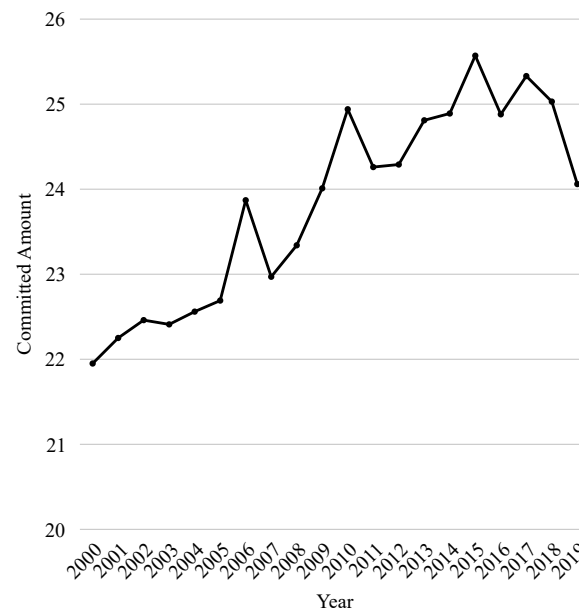
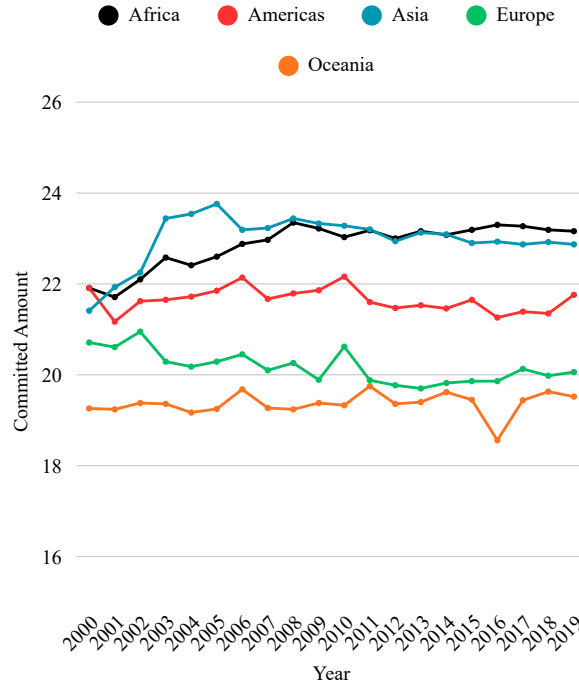


Figure 2: U.S. International Aid

(a) Committed amount of dollars (log) per year and by region



(b) Total committed amount of dollars (log) per year

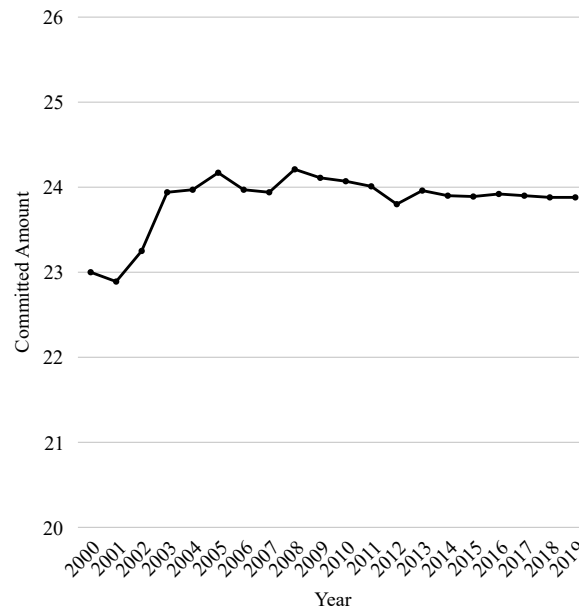
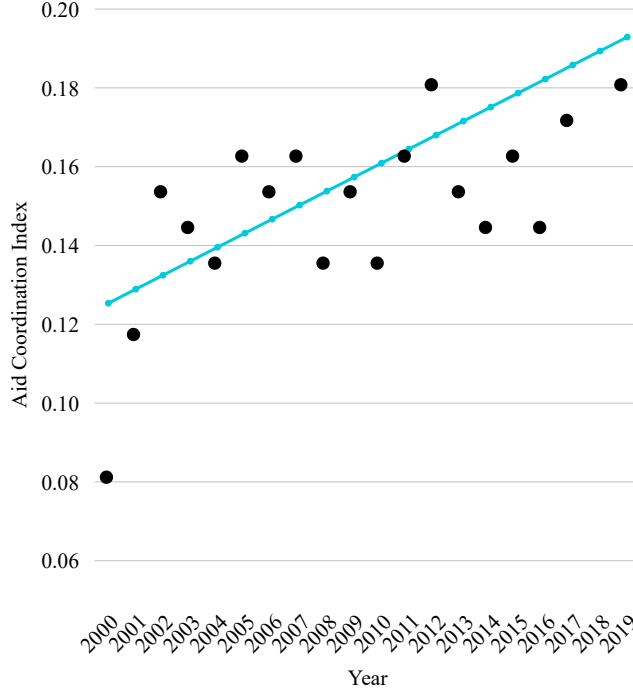


Figure 3: Aid Coordination Index between US and China

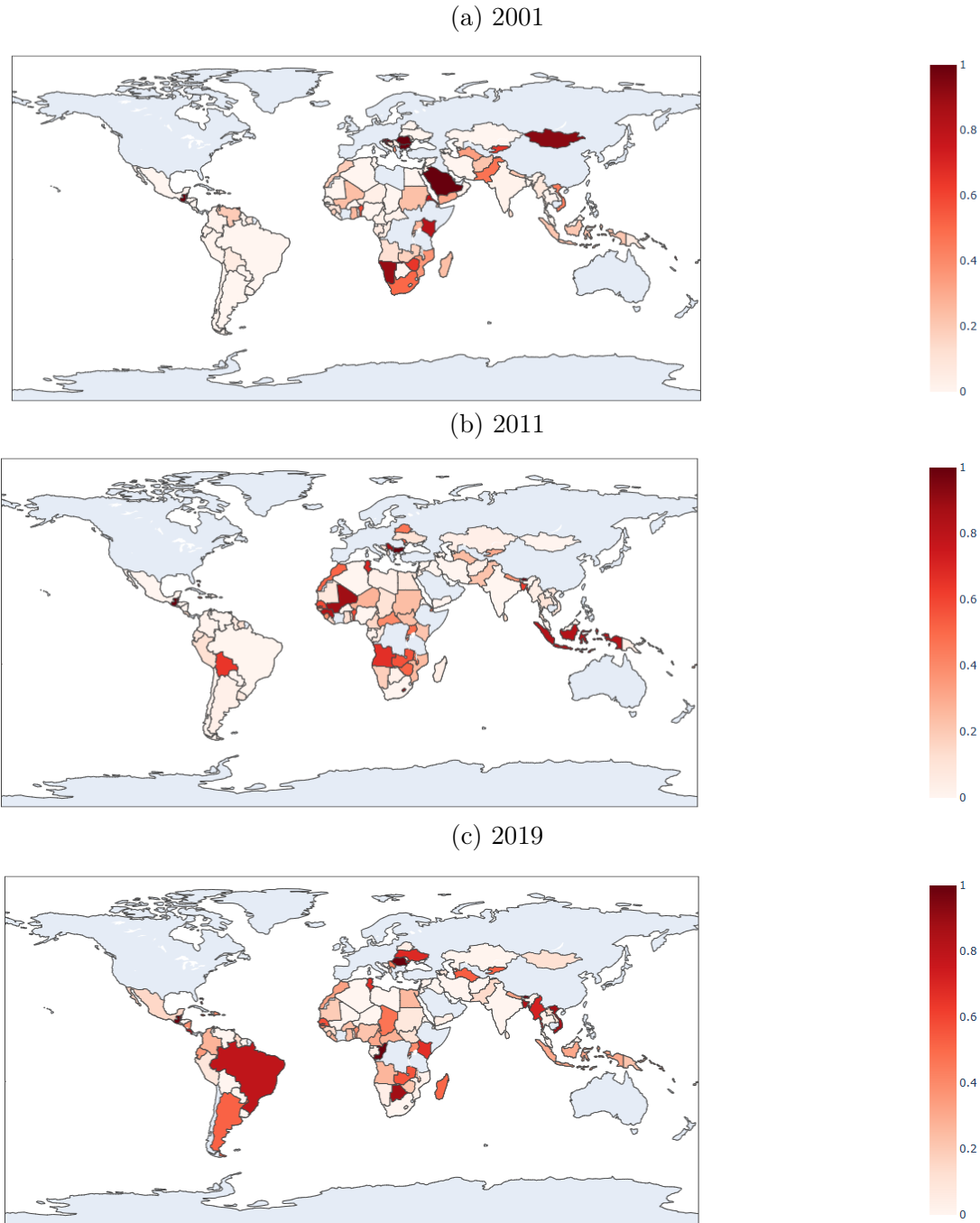


Notes: The graph shows the donors coordination index proposed by [Fuchs et al. \(2015\)](#), and its linear fit, between the U.S. and China over the sample period. The index is computed as follows. For a pair of donors d_1 d_2 , let $x_{1,t}^i$ and $x_{2,t}^i$ be the share of funds that the two donors allocate to a recipient country i in a given period t . The coordination aid index is then computed as $\frac{1}{N} \sum_{i=1}^N \frac{\min\{x_{1,t}^i, x_{2,t}^i\}}{\max\{x_{1,t}^i, x_{2,t}^i\}}$

strategies by using the donor coordination index proposed by [Fuchs et al. \(2015\)](#). The index takes a value of zero when at least one donor is inactive in a recipient country for a given year. Conversely, the index equals one when both donors allocate the same proportion of their total aid budget to the same recipient country in a given year. Figure 3 plots the yearly average coordination index between the U.S. and China over the sample period. The upward trend highlights that, over the years, U.S. and China aid allocations have become increasingly similar. Furthermore, as indicated by Figure 4, the increase in the similarity of China's and the United States' international aid portfolios is mainly attributable to countries in South America and Africa.

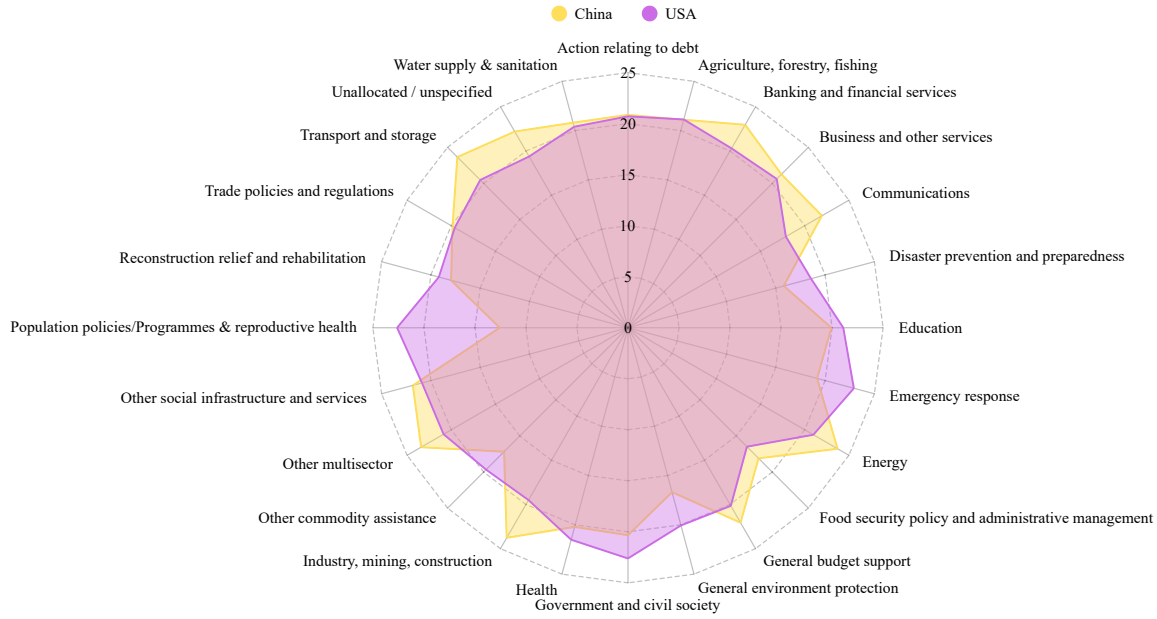
In Figure 5, we report the average amount that China and the U.S. have committed to allocating across various sectors, defined in accordance with the OECD's sector categorization scheme. In general, the U.S. allocation across the various sectors appears to be fairly homogeneous, with a slight prevalence of programs in the health and education sectors. China, instead, concentrates aid more on basic infrastructure and industrial development.

Figure 4: U.S.–China Coordination Index Across Countries



The map shows the coordination index proposed by [Fuchs et al. \(2015\)](#) for U.S. and China in three different years.

Figure 5: Committed Amount in Aid Across Sectors



For details about the sector categorization scheme, see the OECD website.

We conclude this section by presenting a panel regression augmented with year and country fixed effects. The dependent variable is the amount of USD dollars committed by the U.S. to a recipient country in a given year (log scale), and the independent variable is the amount of USD dollars committed by China to a recipient country in a given year (log scale). The regression results are shown in Table 1. There is a positive correlation between Chinese and U.S. aid. However, this correlation disappears once we account for country and time fixed effects, as well as recipient characteristics that are likely to influence the amount of aid a country receives. Specifically, we control for the recipient country's real GDP per capita. The rationale for including the squared term is to account for potential nonlinear effects, as aid is generally expected to be provided when the recipient's income level is neither too high nor too low (Alesina and Dollar, 2000). We also control for the political stability of the recipient country (sourced from the Worldwide Governance Indicators database of the World Bank).

Table 1: Panel Regression

	(1)	(2)	(3)	(4)
	U.S. Aid	U.S. Aid	U.S. Aid	U.S. Aid
China Aid	0.0416** (0.0168)	0.0191 (0.0150)	0.0337** (0.0132)	0.0119 (0.0118)
Real GDP PC			11.89*** (2.582)	9.387* (5.356)
Real GDP PC (square)			-0.794*** (0.155)	-0.396 (0.344)
Political Stability Index			-1.220** (0.533)	-0.891 (0.603)
Constant	13.62*** (0.521)	19.77*** (0.417)	-29.33*** (10.93)	-32.96 (22.83)
Observations	3,240	3,240	1,935	1,935
Number of Country	162	162	138	138
Country FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES
Clustered standard errors in parentheses				

*** p<0.01, ** p<0.05, * p<0.1

3.2 Instrumental variable

The estimates presented in Table 1 do not inform about the causal nexus between the American and Chinese aid. In fact, the Chinese aid allocation criteria are likely to be themselves a function of the American aid allocation. To overcome this issue, we follow [Dreher et al. \(2021\)](#) by instrumenting Chinese financing to a recipient country through a composite term that reflects the overall supply of Chinese funds in a given year and the probability that the country recognizes the government of Taiwan.

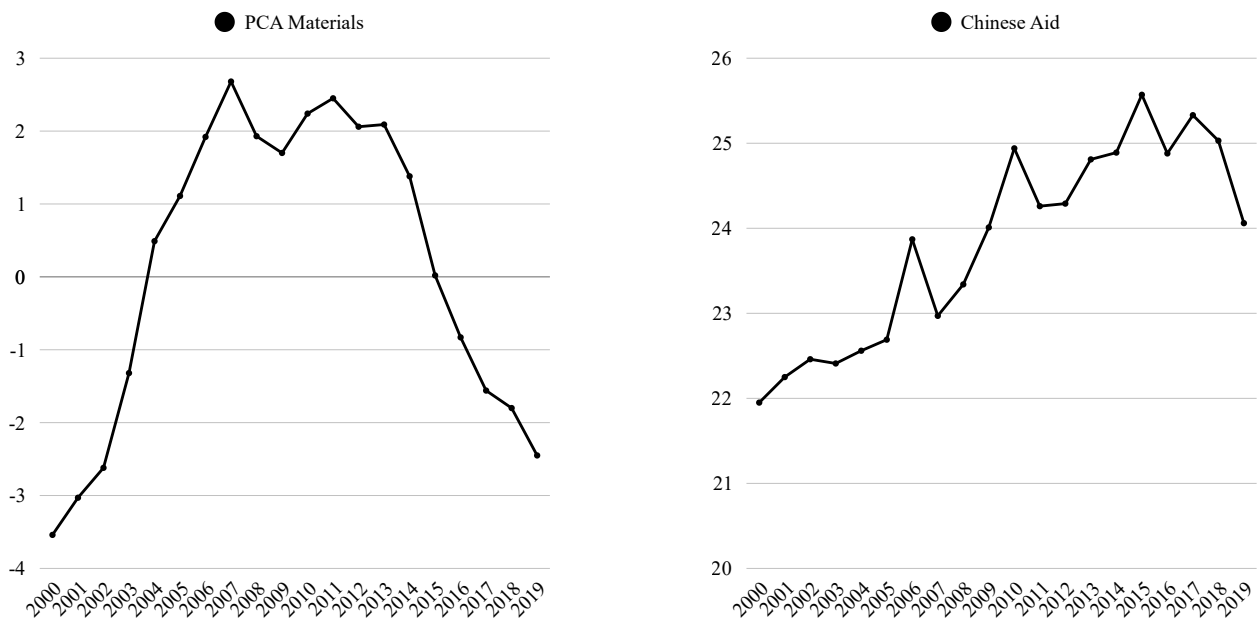
3.2.1 Chinese overproduction

The Chinese government regards domestic inputs such as steel and aluminum as strategic commodities and thus manufactures these materials at levels significantly exceeding domestic consumption. International aid campaigns therefore represent an outlet for this excess production. Thus, there should be a positive correlation between Chinese overproduction and Beijing's willingness to initiate overseas projects.

This reasoning applies broadly across all sectors, as most projects in various domains involve

physical construction and depend heavily on imported Chinese construction materials (Bluhm et al., 2025; Dreher et al., 2021). Therefore, China’s excess production of raw materials is likely a key factor that influences the general level of official Chinese funding. To analyze this, we obtain data on the Chinese production of five raw materials—aluminum, cement, glass, iron, and steel—from the China National Bureau of Statistics. Following the approach of Dreher et al. (2021), we extract the principal component of the cyclical fluctuations within these series.⁷ Figure 6 displays, for each year, the principal component of the cyclical fluctuations in China’s raw material production, alongside China’s total annual commitment (in log) in international aid for the current year.

Figure 6: Chinese overproduction as predictor of Chinese overseas financing



3.2.2 Relationship with Taiwan

Guided by its One China principle, the Chinese government opposes Taiwan from developing diplomatic relations with other countries, or any relations of a state-to-state nature, and opposes Taiwan from participating in the United Nations system and other intergovernmental international organizations (Rich, 2009). In this perspective, the fact that a recipient country does not have a diplomatic relationship with Taiwan can be interpreted as support for the One China principle. Therefore, China should allocate more aid to recipient countries that do not recognize Taiwan Dreher et al. (2021).

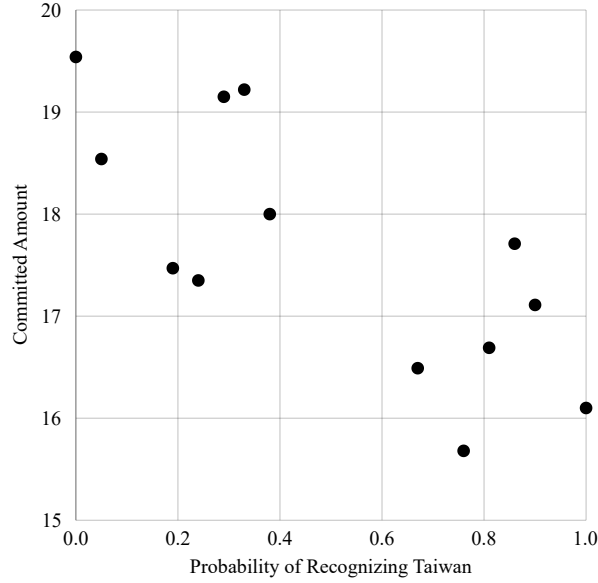
We source data on the political recognition of Taiwan from the Ministry of Foreign Affairs

⁷We extrapolate the cyclical component by applying the Hodrick-Prescott filter to the series.

of Taiwan. For each country-year observation, we create a dummy variable that takes the value of 1 if the country recognizes the government of Taiwan in that year, and 0 otherwise. We then average this dummy variable over the entire sample period. In this way, we obtain a variable capturing the probability that, in a given year, a recipient country recognizes the government of Taiwan. In our sample, for 98 countries this probability is null, for 7 countries it is always equal to 1, and for 10 it is in the interval $(0, 1)$.

Figure 7 plots the average Chinese commitment in international aid (on a log scale) for the different observed values of the probability of recognizing Taiwan in our sample. The decreasing relationship suggests that the longer a recipient country maintains diplomatic relations with Taiwan, the lower the amount of aid it receives from China.

Figure 7: The Recognition of Taiwan as a Predictor of Chinese Aid



3.3 IV regression

In accordance to the above discussion, the following model represents the first stage of our instrumental variable estimation.

$$\text{China Aid}_{i,t} = \beta \times \text{Chinese Overproduction}_{t-1} \times (1 - \text{Prob. Taiwan}_i) + \text{Controls}_{i,t} + \epsilon_{i,t} \quad (1)$$

Where: (i) $\text{China Aid}_{i,t}$ is the amount of USD that China commits to extent to the recipient country i in year t , (ii) $\text{Chinese Overproduction}_{t-1}$ is the principal component of the Chinese inputs overproduction computed as described above, and (iii) Prob. Taiwan_i is the fraction of

years between 2000 and 2019 during which the receiving country i recognizes the government of Taiwan. Note that for the overproduction, we follow [Dreher et al. \(2021\)](#) and use a one-year lag with respect to the aid commitment.

As discussed above, both terms PC_{t-1} and $1 - p_i$ should be positively associated with the Chinese supply of aid. Therefore, we expect a positive value for β . Table 2 column (1) shows the estimates for the first stage. As expected, the instrument we adopt is strongly positively associated with Chinese financing.

Table 2: IV Results

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.830*** (0.225)	
China Aid		0.220* (0.124)
Real GDP PC	-6.296 (5.549)	7.486* (4.051)
Real GDP PC (square)	0.07 (0.344)	-0.287 (0.262)
Political Stability Index	0.545 (0.692)	-0.632 (0.434)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	13.54	
Clustered standard errors in parentheses		

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The second column presents the results for the second stage. When we instrument for Chinese financing, a positive statistically significant relationship with U.S. aid commitments emerges. This suggests that the aid allocation criteria of the two superpowers appear to be strategically complementary.

3.4 Robustness

In this section, we run a bunch of robustness exercises to strengthen the results we stated above.

3.4.1 Additional controls

The results presented in Table 2 remain consistent when we include variables identified by the literature as determinants of aid (Table 3). Specifically, we include a dummy variable that takes the value 1 for country-year observations in which the recipient country is under an IMF arrangement. The rationale behind this control is based on the evidence showing that multilateral organizations, like the IMF, allocate funds based on the political interests of their main stakeholders (Thacker, 1999; Calomiris, 2000; Aiyar et al., 2024). Since the U.S. is a principal player in the IMF, if a strategic interaction between Beijing and Washington is in place, the U.S. could also act through the IMF lending programs. We also include the number of deaths by natural disasters in a given year in a recipient country, as natural disasters may act as an aid catalyst regardless of any strategic interactions between donors. We include variables capturing the political distance between the donors and the recipient countries. To capture political distance, we use the Ideal Point Estimate constructed by Bailey et al. (2017), which is based on vote similarity at the General Assembly of the UN. Finally, we also include trade activity between donors and the recipient country (source: COMTRADE).

Table 3: Robustness: Additional Controls

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.756*** (0.231)	
China Aid		0.3186** (0.159)
Real GDP PC	-5.692 (5.815)	7.574* (4.590)
Real GDP PC (square)	-0.005 (0.360)	-0.227 (0.300)
Political Stability Index	0.545 (0.686)	-0.816* (0.454)
IMF Lending	-0.784* (0.431)	0.874** (0.365)
Death by Natural Disasters	0.547*** (0.017)	-0.014 (0.010)
Political Distance (China)	-0.363 (0.637)	0.396 (0.441)
Political Distance (USA)	-0.207 (0.702)	-0.174 (0.513)
China Export	0.076 (0.072)	0.020 (0.044)
China Import	-0.019 (0.059)	-0.076 (0.052)
U.S. Export	0.179** (0.073)	-0.081* (0.434)
U.S. Import	-0.104* (0.059)	-0.043 (0.052)
Observations	2152	2152
Number of Country	135	135
Country FE	YES	YES
Year FE	YES	YES
F statistics	10.63	

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.4.2 ODA

As we discussed above, in our dataset the vast majority of U.S. projects are categorized as ODA. Instead, around 77% of Chinese projects fall in this category. However, the positive association between U.S. and Chinese aid allocation is still present also when we only consider Chinese financing that can be categorized as ODA (Table 4).

Table 4: Robustness: ODA Aid

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.743*** (0.220)	
China Aid		0.24* (0.139)
Real GDP PC	-5.814 (5.235)	7.531** (3.829)
Real GDP PC (square)	0.084 (0.324)	-0.293 (0.245)
Political Stability Index	0.983 (0.639)	-0.754* (0.434)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	11.35	
Clustered standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

A potential concern regarding our instrument is that the probability of recognizing Taiwan has low cross-sectional variance. In our sample, only for 30 countries is the term $(1-p_i)$ different from 0. For these countries, the cross-sectional variability of the first stage arises solely from the country fixed effects and the controls. As a result, the standard deviation of the term $(1-p_i)$ is 0.33.

To address this concern, we repeat the estimation proposed in Table 2 by adding an additional instrument to the first-stage specification described by (1). Following the same reasoning we adopt for the main instrument, we build the additional instrument by interacting the principal component of the cyclical fluctuations in China's raw material production with country-specific variables capturing the diplomatic and commercial relationship of the recipient country with China.

In Table 5, we use as an additional instrument the interaction between the principal component of the cyclical fluctuations in China's raw material production and the sample average political proximity between the recipient country and China⁸. In Table 6, instead, we use as an additional instrument the interaction between the principal component of the

⁸We compute the political proximity as the inverse of the political distance we adopt in Table 3

cyclical fluctuations in China’s raw material production and the recipient countries’ sample average of imports from China (expressed in billions of USD). Finally, in Table 7, we use as an additional instrument the interaction between the principal component of the cyclical fluctuations in China’s raw material production and the recipient countries’ sample average of exports to China (expressed in billions of USD). In all cases, the estimates are pretty similar to those described in our baseline specification. However, the estimate of the Hansen statistic casts doubt on the joint validity of the instruments.

The political proximity and value of trade should capture the economic and political ties between the recipient country and China. To measure these ties for each recipient country, we extrapolate the principal component of the political proximity between the recipient country and China, (ii) the volume of Chinese exports to the recipient country, and (iii) the volume of Chinese imports from the recipient country. We then average this component over the sample period. Table 8 shows the results. When we use the interaction between the principal component of the cyclical fluctuations in China’s raw material production and the principal component of variables that should capture economic and political ties between China and the recipient country, our estimates are similar to those we proposed in our baseline, and there is no evidence against the joint validity of our instruments.

Table 5: Robustness: Additional Instrument
Political Proximity

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.838*** (0.224)	
Chinese Overproduction $\times$ Political Proximity (China)	0.002*** (0.000)	
China Aid		0.181* (0.097)
Real GDP PC	-6.159 (5.534)	7.242* (3.879)
Real GDP PC (square)	0.068 (0.344)	-0.284 (0.254)
Political Stability Index	0.526 (0.691)	-0.615 (0.425)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	11.013	
Hansen statistics	0.632	
Clustered standard errors in parentheses		

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Robustness: Additional Instrument
Chinese Export

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.894*** (0.225)	
Chinese Overproduction $\times$ Chinese Export	-0.020** (0.008)	
China Aid		0.216* (0.111)
Real GDP PC	-6.257 (5.503)	7.460* (4.027)
Real GDP PC (square)	0.068 (0.342)	-0.287 (0.261)
Political Stability Index	0.468 (0.703)	-0.630 (0.437)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	10.721	
Hansen statistics	0.011**	
Clustered standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 7: Robustness: Additional Instrument
Chinese Import

	(1)	(2)
	First Stage	Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.882*** (0.227)	
Chinese Overproduction $\times$ Chinese Import	-0.016 (0.012)	
China Aid		0.225* (0.120)
Real GDP PC	-6.376 (5.541)	7.521* (4.072)
Real GDP PC (square)	0.078 (0.344)	-0.288 (0.263)
Political Stability Index	0.497 (0.695)	-0.635 (0.439)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	9.806	
Hansen statistics	0.022	
Clustered standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 8: Robustness: Additional Instrument
Principal Component

	(1) First Stage	(2) Second Stage
Chinese Overproduction $\times(1-\text{Prob. Taiwan})$	0.896*** (0.226)	
Chinese Overproduction $\times\text{PCA(Political Proximity (China), Chinese Export, Chinese Import)}$	-0.003** (0.0016)	
China Aid		0.220* (0.120)
Real GDP PC	-6.311 (5.517)	7.486* (4.046)
Real GDP PC (square)	0.073 (0.343)	-0.287 (0.262)
Political Stability Index	0.473 (0.700)	-0.632 (0.438)
Observations	2197	2197
Number of Country	138	138
Country FE	YES	YES
Year FE	YES	YES
F statistics	10.516	
Hansen statistics	0.00***	
Clustered standard errors in parentheses		

*** p<0.01, ** p<0.05, * p<0.1

3.5 Heterogeneity

Table 9: Heterogeneity (Control Function Approach): Trade, Political distance, and time

	(1)	(2)	(3)
	Second Stage	Second Stage	Second Stage
China Aid	0.379*** (0.136)	0.213* (0.115)	0.224* (0.116)
Real GDP PC	7.355** (3.546)	6.979** (3.512)	6.289* (3.741)
Real GDP PC (square)	-0.286 (0.235)	-0.266 (0.233)	-0.216 (0.251)
Political Stability Index	-0.585 (0.430)	-0.660 (0.440)	-0.662 (0.429)
Political Distance (USA)	0.321 (0.452)		
Political Distance (USA)×China Aid	-0.0589** (0.0235)		
American Export		-0.102** (0.0453)	
American Export×China Aid		0.00266* (0.00157)	
American Import		0.0615** (0.0253)	
American Import×China Aid		-0.00176 (0.00107)	
Post Xi Jinping			-2.783** (1.241)
Post Xi Jinping×China Aid			0.0503* (0.0287)
Observations	2,198	2,200	2,200
Number of Country	140	141	141
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Clustered standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

The positive relationship between the amount of aid that China commits to providing to a country and the amount that the U.S. allocates suggests that the two donors are unable to coordinate their humanitarian efforts. One explanation for this complementarity is that aid acts as a means to expand soft power and influence over the recipient country. Therefore, a reasonable interpretation of the results documented in the previous sections is that the U.S. is attempting to counter the rise of China's soft power.

With the data at our disposal, we cannot directly test the validity of our interpretation. However, if it is true that U.S. strategies in allocating humanitarian aid are driven by the

need to limit China’s ascent in soft power, it is reasonable to expect that the U.S. response to Chinese financing is greater when the U.S. has a particular interest in the recipient country. Consequently, we test whether the U.S. is more likely to increase its aid response when it maintains stronger economic or political ties with the recipient country. Since these relationships serve broader strategic objectives beyond humanitarian concerns, if we observe that the U.S. responds more strongly to Chinese aid when it is distributed in countries with strong political or commercial ties to the U.S., we will interpret this as evidence that the motivation behind the American response is rooted in efforts to curb the expansion of Chinese influence.

In Table 9, we explore the interaction between Chinese financing and two factors: the political distance between the recipient country and the U.S. in column (1), and trade (exports and imports) between the recipient country and the U.S. in column (2). The estimates indicate that the U.S. responds more strongly to Chinese aid when American exports rise and when there is greater political proximity between the recipient country and the U.S.

Finally, we explore the role of time in affecting the U.S. response to Chinese aid. As discussed in the introductory section, China has significantly increased its supply of international aid since Xi Jinping’s election, with the manifest goal of expanding Chinese soft power (Shambaugh, 2015; Li, 2022). Therefore, if the U.S. aims to limit the rise of Chinese soft power through its own international aid policy, it is reasonable to expect that the U.S. response to Chinese financing would be stronger from the moment China demonstrates its intention to define its own sphere of influence. In column (3) of Table 9, we interact Chinese financing with a dummy variable that takes the value of 1 for years following Xi Jinping’s election (2013), and 0 otherwise. The coefficient associated with this interaction term is positive, which we interpret as a hint that the U.S. begins to respond more strongly to Chinese aid when China starts to prioritize the expansion of soft power.

4 Conclusion

In this paper, we investigated the interplay between U.S. and Chinese international aid strategies, framing it as a potential arena for strategic competition. Our analysis reveals a positive association between Chinese and American aid financing, suggesting that the U.S. tends to increase its aid commitment in countries where China also commits more funds. This finding points to a dynamic where the U.S. may be attempting to counter the rise of China’s soft power by strategically allocating its aid in response to Chinese financing. Our results highlight the complex dynamics of international aid, suggesting that it is not solely driven by altruistic motives but also by strategic considerations related to great power competition. The competition between the U.S. and China in establishing their spheres of influence may lead to aid fragmentation, potentially undermining the effectiveness of aid and the stability of recipient countries.

References

- Aiyar, S., D. Malacrino, and A. F. Presbitero (2024). Investing in friends: The role of geopolitical alignment in fdi flows. *European Journal of Political Economy* 83, 102508.
- Alesina, A. and D. Dollar (2000). Who gives foreign aid to whom and why? *Journal of economic growth* 5, 33–63.
- Alesina, A. and B. Weder (2002). Do corrupt governments receive less foreign aid? *American economic review* 92(4), 1126–1137.
- Allison, G. (2017). The thucydides trap. *Foreign Policy* 9(6), 73–80.
- Annen, K. and L. Moers (2017). Donor competition for aid impact, and aid fragmentation. *The World Bank Economic Review* 31(3), 708–729.
- Bailey, M. A., A. Strezhnev, and E. Voeten (2017). Estimating dynamic state preferences from united nations voting data. *Journal of Conflict Resolution* 61(2), 430–456.
- Bauer, P. T. (1976). *Dissent on development*. Harvard University Press.
- Bermeo, S. B. (2011). Foreign aid and regime change: A role for donor intent. *World Development* 39(11), 2021–2031.
- Berthélemy, J. et al. (2006). Aid allocation: Comparing donor’s behaviours. *Swedish economic policy review* 13(2), 75.
- Blair, R. A., R. Marty, and P. Roessler (2022). Foreign aid and soft power: Great power competition in africa in the early twenty-first century. *British Journal of Political Science* 52(3), 1355–1376.
- Bluhm, R., A. Dreher, A. Fuchs, B. C. Parks, A. M. Strange, and M. J. Tierney (2025). Connective financing: Chinese infrastructure projects and the diffusion of economic activity in developing countries. *Journal of Urban Economics* 145, 103730.
- Boone, P. (1996). Politics and the effectiveness of foreign aid. *European economic review* 40(2), 289–329.
- Burnside, C. and D. Dollar (2000). Aid, policies, and growth. *American economic review* 90(4), 847–868.
- Calomiris, C. W. (2000). When will economics guide, imf and world bank reforms. *Cato J.* 20, 85.
- Chung, E., A. O. Pechenkina, and K. K. Skinner (2023). Competitors in aid: how international rivalry affects public support for aid under various frames. *Political Research Quarterly* 76(3), 1371–1387.

- Dasandi, N. and L. Erez (2019). The donor’s dilemma: International aid and human rights violations. *British Journal of Political Science* 49(4), 1431–1452.
- Djankov, S., J. G. Montalvo, and M. Reynal-Querol (2009). Aid with multiple personalities. *Journal of Comparative Economics* 37(2), 217–229.
- Dreher, A., A. Fuchs, R. Hodler, B. C. Parks, P. A. Raschky, and M. J. Tierney (2021). Is favoritism a threat to chinese aid effectiveness? a subnational analysis of chinese development projects. *World Development* 139, 105291.
- Dreher, A., A. Fuchs, B. Parks, A. Strange, and M. J. Tierney (2021). Aid, china, and growth: Evidence from a new global development finance dataset. *American Economic Journal: Economic Policy* 13(2), 135–174.
- Dreher, A., P. Nunnenkamp, and M. Schmaljohann (2015). The allocation of german aid: Self-interest and government ideology. *Economics & Politics* 27(1), 160–184.
- Dudley, L. and C. Montmarquette (1976). A model of the supply of bilateral foreign aid. *The American Economic Review* 66(1), 132–142.
- Easterly, W. (2003). Can foreign aid buy growth? *Journal of economic Perspectives* 17(3), 23–48.
- Easterly, W. (2007). Are aid agencies improving? *Economic policy* 22(52), 634–678.
- Easterly, W. (2009). Can the west save africa? *Journal of economic literature* 47(2), 373–447.
- Easterly, W., R. Levine, and D. Roodman (2004). Aid, policies, and growth: comment. *American economic review* 94(3), 774–780.
- Edel, C. and H. Brands (2019). The real origins of the us-china cold war. *Foreign Policy* 2.
- Fajgelbaum, P., P. Goldberg, P. Kennedy, A. Khandelwal, and D. Taglioni (2024). The us-china trade war and global reallocations. *American Economic Review: Insights* 6(2), 295–312.
- Fuchs, A., P. Nunnenkamp, and H. Öhler (2015). Why donors of foreign aid do not coordinate: The role of competition for export markets and political support. *The World Economy* 38(2), 255–285.
- Goodman, S., S. Zhang, A. A. Malik, B. C. Parks, and J. Hall (2024). Aiddata’s geospatial global chinese development finance dataset. *Scientific Data* 11(1), 529.
- Hopewell, K. (2020). *Clash of powers: US-China rivalry in global trade governance*. Cambridge University Press.

- Kennedy, A. B. and D. J. Lim (2018). The innovation imperative: technology and us–china rivalry in the twenty-first century. *International Affairs* 94(3), 553–572.
- Kersting, E. and C. Kilby (2014). Aid and democracy redux. *European Economic Review* 67, 125–143.
- Knack, S. (2004). Does foreign aid promote democracy? *International studies quarterly* 48(1), 251–266.
- Knack, S. and A. Rahman (2007). Donor fragmentation and bureaucratic quality in aid recipients. *Journal of development economics* 83(1), 176–197.
- Lang, V. F. and A. F. Presbitero (2018). Room for discretion? biased decision-making in international financial institutions. *Journal of Development Economics* 130, 1–16.
- Lebovic, J. H. (2005). Donor positioning: development assistance from the us, japan, france, germany, and britain. *Political Research Quarterly* 58(1), 119–126.
- Li, J. (2022). China under xi jinping. *Journal of International Affairs* 75(1), 261–272.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy¹. *American political science review* 53(1), 69–105.
- McFaul, M. (2020). Cold war lessons and fallacies for us-china relations today. *The Washington Quarterly* 43(4), 7–39.
- Morgenthau, H. (1962). A political theory of foreign aid. *American political science review* 56(2), 301–309.
- Morgenthau, H. J., K. W. Thompson, and W. D. Clinton (1985). *Politics among nations: The struggle for power and peace*, Volume 688. Knopf New York.
- Natsios, A. S. (2020). Foreign aid in an era of great power competition. *Prism* 8(4), 100–119.
- O’Keefe, J. (2007). Aid: From consensus to competition. *Global Impact: Philanthropy Changing Development*, *Brookings Blum Roundtable Agenda*, August 3.
- Rich, T. S. (2009). Status for sale: Taiwan and the competition for diplomatic recognition. *Issues & Studies* 45(4), 159–188.
- Rosenstein-Rodan, P. N. (1961). International aid for underdeveloped countries. *The review of economics and statistics*, 107–138.
- Shambaugh, D. (2015). China’s soft-power push: The search for respect. *Foreign Affairs* 94(4), 99–107.

- Steinwand, M. C. (2015). Compete or coordinate? aid fragmentation and lead donorship. *International Organization* 69(2), 443–472.
- Svensson, J. (1999). Aid, growth and democracy. *Economics & politics* 11(3), 275–297.
- Thacker, S. C. (1999). The high politics of imf lending. *World politics* 52(1), 38–75.
- Tierney, D. (2024). Forever competition: the end-game of sino–us rivalry. *International Affairs* 100(1), 361–383.
- Wellner, L., A. Dreher, A. Fuchs, B. C. Parks, and A. Strange (2025). Can aid buy foreign public support? evidence from chinese development finance. *Economic Development and Cultural Change* 73(2), 523–578.
- Whang, T., Y. Kim, J. T. Han, and H. J. Kim (2019). Us foreign aid and economic policy concessions. *Policy Studies* 40(1), 58–79.
- Younas, J. (2008). Motivation for bilateral aid allocation: Altruism or trade benefits. *European journal of political economy* 24(3), 661–674.

Authority, Information, and Credit Terms: Evidence from Small Business Lending

Andrea Bellucci

Department of Economics, University of Insubria and MoFiR

`andrea.bellucci@uninsubria.it`

Alexander Borisov

Carl H. Lindner College of Business, University of Cincinnati and MoFiR

`alexander.borisov@uc.edu`

Alberto Zazzaro

Department of Economics and Statistics, University of Naples Federico II, CSEF, and MoFiR

`alberto.zazzaro@unina.it`

Abstract

This paper empirically examines how the allocation of decision-making authority and the availability of information affect credit terms in small business lending. Using a sample of credit lines and variation arising from the organizational restructuring strategy of an Italian bank, we show that an increase in the overlap between decision-making authority and information production reduces the use of collateral but leaves interest rates unchanged. The effect is more pronounced when loan officers have access to soft information or tacit local knowledge about borrowers or when the real authority of the loan officers is limited. Our results highlight the interplay between decision-making authority and information production in the context of loan contracting.

Keywords: Soft and hard information, Collateral, Interest rate, Organizational hierarchies, SMEs financing

JEL: D83, D21, G21, G30, L11

1 Introduction

Theory argues that the allocation of decision-making authority within firms and organizations can shape agents' incentives to produce, use, and communicate information (Aghion and Tirole, 1997; Dessein, 2002; Stein, 2002). In the context of bank lending, empirical work shows that the delegation of decision-making authority affects the production of information and the quality and type of information that local loan officers produce and channel to senior bank managers at higher hierarchical levels of the organization. Specifically, granting loan approval authority to the loan officers responsible for information collection and risk analysis motivates production of more and unbiased soft information and fosters a greater effort to manage the lending process with a positive effect on credit allocation, loan performance, and bank profitability (Agarwal and Hauswald, 2010; Calcagnini et al., 2018; Liberti, 2018; Bouwens and Kroos, 2019; Skrastins and Vig, 2019). However, the contract-term implications of delegating decision-making authority to the agents who process and manage loans and collect and evaluate information about (potential) borrowers are relatively underexplored and poorly understood. Hence, in this paper, we aim to fill this gap by examining, to the best of our knowledge, for the first time in the literature, how the allocation of decision-making authority to local loan officers affects collateral requirements (and interest rates) in small business lending.

We start with the notion that collateral arises as a solution to information-asymmetry frictions between lenders and borrowers (Chan and Kanatas, 1985; Besanko and Thakor, 1987; Chan and Thakor, 1987; Boot et al., 1991; Boot and Thakor, 1994). Consequently, a greater effort to produce information through screening and monitoring of borrowers, which reduces informational frictions, should diminish the relevance of mechanisms such as collateral intended to mitigate the effects of these frictions on credit risk. However, how changes in available information are ultimately reflected in collateral requirements and other contract terms depends on decision-making authority. More information, decoupled from the authority to act on it, might not have an impact. In this paper, we use a quasi-experimental empirical design and the context of small business lending to study the implications of variation in the overlap between decision-making authority and information production for collateral requirements.

A challenge for studies on the effects of allocating decision-making authority is that identifying causal effects is difficult due to concerns about endogeneity, self-selection, and measurement error. Relatedly, detailed micro-level data on authority allocation are limited (Liberti, 2018; Skrastins and Vig, 2019). To mitigate these concerns, we adopt an identification approach that explores an organizational change implemented by a medium-sized

regional Italian bank, hereafter *the bank*. The change shifted decision-making authority for some credit contracts from senior managers at the bank headquarters to local loan officers at bank branches, increasing the overlap between information production and decision-making authority. In our study, we trace the effect of this increase.

Specifically, the bank has a lending policy that uses an assignment process, mainly organized around loan-size-based thresholds, to determine whether decisions about a given contract are made at the local bank branch, hereafter *local level*, or at the bank headquarters, hereafter *central level*. The underlying premise of our empirical strategy is that a contract decided and managed at the local level is within the formal authority of the local loan officer, who also produces the relevant information, both hard and soft. By contrast, the decision-making authority for a contract managed at the central level is removed from the locus of information production. This creates friction because the transfer of information within the banking organization inevitably reduces its ability to provide an unbiased signal of the applicant’s risk, especially when the information is soft (Stein, 2002; Alessandrini et al., 2009; Liberti and Petersen, 2019). Stated differently, for local-level contracts, the delegation of decision-making authority to local loan officers is characterized by a complete overlap between the point of information production and its use within the bank. By contrast, if decision-making authority remains at the central level, as with central-level contracts, the overlap is at best partial.

Within this context, we exploit an exogenous shock to the allocation of decision-making authority for certain contracts, resulting from a restructuring strategy implemented by the bank as part of a broader organizational change in 2005. In particular, the size-based thresholds used to determine whether contract decisions are managed locally were increased for some branches. As a result, some contracts that had originated at these branches but had previously been managed at the central level, because they exceeded the pre-restructuring assignment threshold, were returned to the local level. This led to an increase (“shock”) in the decision-making authority of the information-producing agent for some (“treated”) credit lines, while others (“controls”) remained unaffected. Hence, we conduct a difference-in-differences (DD) analysis of the resultant increase in the overlap between available information and decision-making authority to capture its effect on collateral requirements (and interest rates).

For this analysis, we use data on the portfolio of outstanding credit lines granted by the bank to small and medium-sized enterprises (SMEs) in two Italian provinces as of September 2004 and 2006, i.e., before and after the organizational change in 2005. The dataset covers a broad group of borrowers operating across various economic sectors and served by the bank’s branch network in these two provinces. It also contains information on contract terms and

borrower characteristics. Importantly, it indicates the branch where a credit line originated. This allows us to identify credit lines subject to the “shock” and implement our empirical analysis.¹

Our DD estimates indicate that delegating authority to the loan-originating officer reduces the likelihood that the loan is collateralized. The estimated effect is economically important. Based on our main specification, we infer that the incidence of collateral decreases by about 15% for treated contracts, which is substantial given that the unconditional incidence of collateral for these contracts is 40%. In addition to being statistically significant and economically important, the estimate is robust to alternative estimation techniques, controls for the nature of bank-borrower interaction, borrower characteristics, and credit market conditions, and to other robustness tests. We also demonstrate that our results hold if we use as a benchmark an alternative control group of (“untreated”) contracts, i.e., credit lines that are not subject to the implemented restructuring change.

The estimated baseline relationship between collateral and allocation of decision-making authority is consistent with the argument that placing both the production of information and the authority to act on it in the hands of the same loan officer improves information availability and utilization. This reduces reliance on costly contractual instruments, such as collateral, intended to mitigate information frictions between borrowers and lenders. Given that such frictions are central to our arguments, we conduct additional tests to demonstrate that the estimated effect is indeed a realization of increased use of soft information by local loan officers incentivized by the delegation of formal loan approval authority. To this end, we explore variation in the amount of information available about a borrower. Specifically, we focus on the potential divergence between the information sets of the local and central levels, which emerges in the presence of soft information and tacit local knowledge. First, we use the idea that banks accumulate information over the course of a lending relationship, but some of this information is soft and remains at the local branch where the loan is originated (Hattori et al., 2015). Second, we rely on the notion that local loan officers generate and retain tacit, unverifiable knowledge about the local credit market that may be difficult to

¹While establishing external validity of our analysis is not trivial, we point to several aspects of our study that speak to the generalizability of our insights. First, we note that advances in information and communication technology (ICT) and credit scoring methods have contributed to (partial) hardening of soft information over the last couple of decades (Liberti and Petersen, 2019). This has led to improved transmission of such information within banking organizations, thereby favoring the delegation of decision-making authority to lower hierarchical levels. Consistently, such a shift in authority has occurred at many Italian banks, especially small and medium-sized ones, in recent years (Mocetti et al., 2017). Thus, the “shock” to decision-making authority we analyze is generally applicable to a wide range of banks. Second, the context and test setting we explore are fairly representative of medium-sized banks and bank-based economies. We discuss in Appendix B and Section 2.1 how characteristics of the test setting contribute to the external validity of our inferences.

transmit, and that this knowledge increases with their geographic proximity to borrowers.

To the extent that the officers at the lending branch are the repository of soft information and tacit local knowledge, the potential divergence between locally available information and centrally available information is greater for borrowers with long-standing lending relationships or those located close to the lending branch. Consistently, we find that the estimated effect of increased authority on collateralization is stronger when the bank-borrower relationship is longer and when the branch-borrower distance is shorter. The result aligns with the idea that the allocation of decision-making authority is relatively more important when local loan officers are more likely to possess information than senior managers at higher hierarchical levels in the bank (Dessein, 2002; Alonso et al., 2008).

We also explore whether the effect of delegating formal authority to the local loan officers depends on their real authority prior to the organizational change. While senior managers at the bank headquarters may have formal authority, their real authority can be limited if their decisions are based on recommendations and information from local loan officers, and their control over these officers is also limited (Aghion and Tirole, 2007). In other words, if the local loan officers have significant real authority before the restructuring, post-restructuring allocation of formal authority might not result in any significant changes, as the contract terms already reflect the real authority of the local loan officer. By contrast, if the real authority of the loan officers is limited, granting them formal authority can have an effect. We capture the real authority of the local loan officers using branch-headquarters distance. The underlying assumption is that when the branch originating the contract is located far from the bank headquarters and, respectively, senior managers, it is more arduous and costly to verify the accuracy of information transmitted to the headquarters and possibly reverse recommendations made by the local loan officers (Agarwal and Hauswald, 2010; Filomeni et al., 2021). Regression results indicate that the delegation of formal authority is more likely to reduce the use of collateral when the lending branch is close to the bank headquarters, i.e., when the real authority of the loan officers is limited. By contrast, when the local loan officers already had real authority over the lending process, the allocation of formal authority seems less relevant. This suggests that the two types of authority can be substitutes.

While our main analysis focuses on collateral, as a contractual mechanism intended to resolve information frictions between borrowers and lenders, we also analyze interest rates. In this case, however, the effect of delegation of authority is ambiguous. First, the bank's interest rate policy can be set centrally, leaving little discretion to loan officers at local branches over the rates they can charge for specific loans, regardless of the approval authority (Skranstis and Vig, 2019). Second, more information about the risk profile of a loan application can lead to a decrease or increase in the price of credit, depending on the direc-

tion of the additional information and whether loan officers without formal decision-making authority unbiasedly report, especially unverifiable information, to senior managers at the headquarters responsible for the loan (Bouwens and de Kok, 2021; Bouwens and Kroos, 2019).²

We find that, on average, the “treated” contracts have interest rates that are not significantly different from the rates of the untreated contracts in the control group. This result is also independent of the magnitude of the potential divergence between the information sets at the local and central levels. However, we show a significant interaction between real and formal authority. Specifically, we find that delegating formal authority to loan officers with real authority is associated with lower interest rates. Conversely, loan officers with limited real authority before the organizational change ask for higher rates once they have formal authority, suggesting possible distortions in the transmission of information across the bank’s hierarchical levels.

Our work adds to multiple streams of research. Related studies examine how delegation of authority over lending decisions affects credit availability and loan size (Liberti and Mian, 2009; Canales and Nanda, 2012; Skranstis and Vig, 2019; Filomeni et al., 2021; Porzio et al., 2020) and others analyze interest rates (Qian et al., 2015; Liberti, 2018; Bouwens and Kroos, 2019; Bouwens and de Kok, 2021). We are the first to focus on collateral requirements and complement extant research in two additional dimensions. First, we examine lending contracts from a medium-sized regional bank. This departure from the large-bank paradigm, which the (empirical and theoretical) literature on delegation of authority has focused on, highlights the importance of allocating decision-making authority, even when organizational frictions are less pronounced. Second, we explore the interaction between formal and real authority. Agarwal and Hauswald (2010) and Liberti (2018) are the only studies, to our knowledge, that analyze the role of loan officers’ real authority in the lending process. However, they do not explore how the effects of the delegation of formal authority depend on the real authority that loan officers can effectively exert over the lending process and senior bank managers’ decisions.

We also add to the literature on collateral. Theory argues that collateral arises as a solution to problems of information asymmetry. Empirical tests rely on proxies for information availability such as length and exclusivity of the lending relationship (Petersen and Rajan, 1994; Berger and Udell, 1995; Degryse and Van Cayseele, 2000; Elsas and Krahnen, 2000), use of ratings or credit scoring technology (Machauer and Weber, 1998; Berger et

²We note that the change in interest rate is related to the direction (positive or negative) of the information, while the use of collateral is driven by the magnitude of information asymmetry. With more information on approved applications, even if negative, banks can focus on interest rate as long as the use of collateral is associated with additional costs for the bank.

al., 2011), proximity to borrowers or presence in local credit markets (Bellucci et al., 2019), and organizational structure or hierarchical distance (Jimenez et al., 2006; Jimenez et al., 2009). We complement this research by tracing the effect of a change, generated by a largely exogenous shock, in the amount of available information to the party with decision-making authority on collateral.

The rest of the paper is organized as follows. The next section describes data and context and provides details about the identification strategy. Our main results, and some robustness tests, on collateral are in Section 3. In Section 4, we explore underlying mechanisms by detecting factors that moderate the baseline relationship. In Section 5, we examine interest rates as another contract term. We conclude in Section 6.

2 Data and identification strategy

2.1 Context

For our analysis, we utilize a proprietary dataset of credit lines to SMEs by a regional Italian bank. The dataset covers the entire portfolio of outstanding credit lines, as of September 2004 and 2006, of the bank in two provinces in the Marche region of Italy. The bank’s regional headquarters is also located in one of the provinces, and its operations there are carried out through a network of more than 50 branches. We provide more details on the economic context pertaining to the provinces, the structure of the banking sector, and the representativeness of the sample in Appendix B.

Each branch is managed by a branch head, who supervises the loan officers at the branch and reports to headquarters. The tasks of the local loan officers include typical commercial banking activities such as collection and processing of quantitative and qualitative information about applicants, verification of documentation, analysis of creditworthiness and risk of applicants, application approval and setting of contract terms (interest rate, collateral, quantity), management of the lending relationship and periodic renegotiation of contract terms, monitoring, collateral valuation, and reporting to the upper hierarchical levels of the bank.

Under the bank’s lending policy, officers at the local branch can approve loan contracts only if the application amount is below a prespecified threshold. If the threshold is exceeded, the local loan officer transfers the collected information and documentation to upper decisional levels at the bank headquarters, where senior managers make a determination. If the threshold is not exceeded, and the credit line is originated and managed at the branch level, the local loan officer has the authority to set initial contract conditions for both price and

non-price terms, as well as to renegotiate the contract periodically.

While local loan officers have the authority to determine credit conditions in principle, their autonomy in setting contract terms varies by term. For interest rates, for instance, autonomy is limited by criteria and system-wide policies established by the bank and related to factors such as the size of the credit line, the financial condition of the applicant, the internal credit score, etc. Within the policy, loan officers have limited authority to adjust the interest rate by modifying the spread they charge over the base rate. However, their authority and ability to determine collateral requirements are more substantial. Thus, these can reflect the loan officers' information about the borrower, as well as their incentives. The incentive structure of employment contracts of individual loan officers is not disclosed by the bank, but we are aware that branch-level evaluations are based on the amount of credit allocated and performance. Loan officers receive an annual monetary reward linked primarily to the branch's performance, but not to their personal activity within the branch.³

2.2 Identification strategy

The bank's credit-granting process begins at the local branch when the prospective borrower applies for credit. As a result, each credit line originates at one of the local branches. Based on borrower characteristics, the credit line is categorized within the corporate or small-business segments. Contracts in the corporate segment are assigned an account manager whose interaction with the borrower is more comprehensive and direct (e.g., the account manager can visit the borrower or schedule a meeting at the borrower's location). This jeopardizes the key premise of our identification strategy; therefore, we focus only on credit lines in the small-business segment.

While each credit line originates at a local branch of the bank, the bank uses predetermined size-based thresholds to determine if the credit line should be negotiated and managed at the local branch (local level) by the local loan officer. In this case, contract terms and their periodic renegotiations are set and handled at the local branch. If the amount exceeds the threshold, management of the credit line is forwarded to the regional headquarters of the bank (central level) along with all documentation and information collected by the local loan officer. Contract terms and conditions are then determined by the senior management at the headquarters. This transition is central to the identification strategy.⁴

³The incentive scheme is consistent with common incentive structures of Italian banks at the time. A survey conducted by the Bank of Italy on more than 300 banks indicates that in 90% of cases, the overall profitability of the local branch is the primary basis for providing incentives to loan officers (Albareto et al., 2011).

⁴We note that in our dataset we observe only approved credit lines and have no information on applications/rejections.

A key distinction between decisions taken at the local level and those taken centrally is in the information available to the decision-maker. During the contracting process, bank officers interact with borrowers and generate a substantial amount of information. Some of it is “hard” in nature. It can be measured, quantified, and easily communicated with others (e.g., balance sheet data, financial ratios, past default). Other information is “soft”. It is non-transferable and non-verifiable (Liberti and Petersen, 2019). Hence, soft information remains primarily with the agent who produces and collects it, i.e., the local loan officer. While some soft information can be “hardened” when forwarded from the local to the central level, some is lost in the process. Thus, decisions made at a local level can reflect both soft and hard information, while those made at the central level reflect mostly the latter.

In 2005, around the midpoint of our sample period, the bank implemented a restructuring initiative. One of the objectives of the initiative was to reduce the number of internal decision-making layers to shorten communication times, facilitate information transfer within the bank’s organizational hierarchy, and ultimately speed up decision-making. As a part of the restructuring, the size-based thresholds for local decision-making were modified at 5 branches. We refer to them as “treated” branches. The thresholds for the remaining branches were unchanged at their pre-restructuring levels. We note that the “treated” branches were slightly larger. The average number of loan officers at these branches is 3.8, compared with 2.5 at the other local bank branches.

The restructuring initiative doubled the thresholds for local decision-making at the “treated” branches. Specifically, before the restructuring, the threshold at each branch of the bank was €150,000. For credit lines exceeding this threshold, all documentation and information are transferred to the upper decision-making level at the bank headquarters. After the restructuring, branches not affected by the change retained their thresholds, while branches subject to the change had their limit increased to €300,000.⁵ The change was implemented at only some of the branches. As a result, some credit lines originated at the treated branches that were above the local management threshold in 2004 and were therefore managed centrally were returned to the local level in 2005, i.e., post-restructuring. We identify 48 such contracts and denote them by “treated” credit lines. We illustrate the core of our identification strategy in Figure 1.

We denote by C central-level decision-making and by L local-level decision-making. On the left panel of the figure, we depict the “transition” of treated credit lines from the central level pre-restructuring to the local level post-restructuring at treated branches. The number

⁵These figures are consistent with decision-making thresholds of smaller Italian banks. Albareto et al. (2011) find that, on average, the amount of credit loan officers at such banks can approve on their own is up to €200,000.

in parentheses identifies the number of contracts in the respective group.

We also construct a benchmark (or control) group used to approximate the counterfactual. Specifically, the control group includes all comparable credit lines, in terms of size, that are not treated. In other words, we identify credit lines that fall within the treatment range but are not from a treated branch. Since the thresholds are size-based, the control group removes the effect of factors related to loan size but not included in our empirical specification. We have 220 such contracts. In Figure 1, the control group is in the middle panel and is denoted by Control 1. In the control group, we effectively match on size central-level credit lines as of the pre-restructuring period. As the panel illustrates, these are credit lines that remain at the central, or *C*, level after the restructuring. By construction, these credit lines are from non-treated branches.

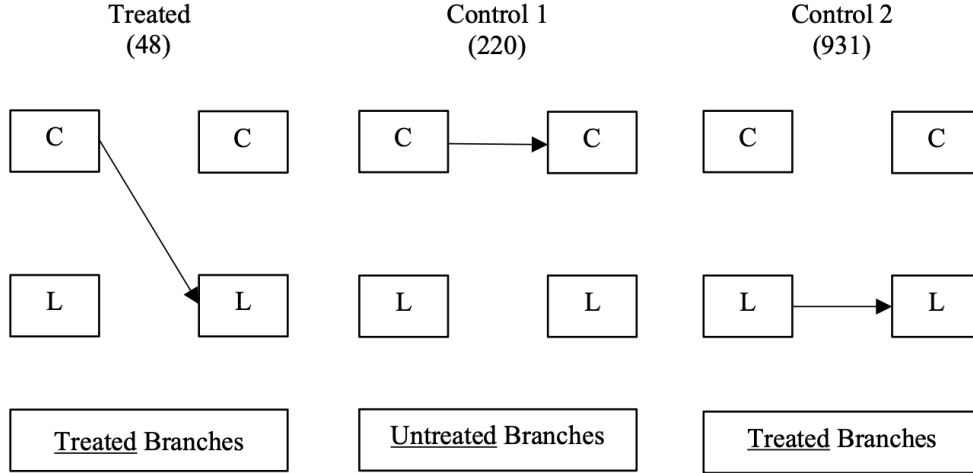


Figure 1: Illustration of the identification strategy.

The figure illustrates our identification approach. The group of treated contracts is denoted by Treated, and the two control groups are denoted by Control 1 and Control 2. Central-level decision making that takes place at bank headquarters is denoted by *C*. Local-level decision making that takes place at local branches is denoted by *L*. The restructuring strategy implemented by the bank doubled the thresholds for local-level decision-making at 5 treated” branches. As a result, some contracts that were managed centrally pre-restructuring are transferred to the local level post-restructuring. This transfer from the central level *C* to the local level *L* is illustrated in the left panel. There are 48 contracts in the Treated group. We construct a benchmark using contracts not subject to treatment. The first control group, Control 1, is illustrated on the central panel of the figure. Control 1 consists of credit lines that are similar in size and are not treated. These are credit lines that fall within the treatment range but are from branches not subject to restructuring, i.e., untreated branches. Hence, they are at central level *C* pre-restructuring and remain at central level *C* post-restructuring. There are 220 contracts in Control 1. We construct another control group, Control 2, shown on the right panel of the figure. Control 2 consists of all credit lines managed locally in both years and originating at the treated branches. These credit lines are at the local level *L* pre-restructuring and remain at the local level *L* post-restructuring. Control 2 consists of 931 contracts. Our main analysis uses Treated and Control 1, while Treated and Control 2 are used for robustness tests.

A key challenge with the above benchmark is that, as mentioned, credit lines in the control

group come from untreated branches. While we account for all time-invariant branch-specific factors in some of our estimates via branch fixed effects, in a robustness test, we construct an alternative control group, Control 2. This group, comprising 931 credit lines, focuses only on the treated branches but relaxes the size requirement. Specifically, we use all credit lines managed locally in both years and originating from the treated branches. Thus, we remove the effects of factors specific to the bank’s branches undergoing restructuring. We also control for credit line amount in these estimations to ensure size differences are not driving the results. This control group is illustrated on the right panel of Figure 1.

Last, given that we use outstanding credit lines for the analysis, our identification strategy assumes that these were renegotiated after the reorganization. While we do not know when each credit line is renegotiated during the sample period, we note that more than 85% of the credit lines we study show a change in credit limit, and more than 95% show a change in interest rate, from 2004 to 2006. For the entire sample of credit lines, these figures are about 81% and 99%, respectively. This supports the underlying assumption of our analysis.

2.3 Empirical model and variables

The restructuring process implemented at the bank discussed above leads to an arguably exogenous shock that increases the alignment of decision-making authority and information production. Using the shock, we adopt a DD estimation strategy to capture the effect of the contracting process on outcomes. Hence, we focus on borrowers present in both 2004 and 2006, and estimate the empirical specification outlined in equation (1) below:

$$Y_{i,t} = \beta_0 + \beta_1 D(Post)_t + \beta_2 D(Treated)_i + \beta_3 D(Treated)_i \times D(Post)_t + \gamma X_{i,t} + \varepsilon_{i,t} \quad (1)$$

In the specification, i denotes borrower and t denotes time period, while Y is a contract term of interest (collateral or interest rate). We estimate our models using ordinary least squares (OLS) in the main part of the analysis, but we verify the robustness of our results to alternative methods that better reflect the nature of some outcome variables. To account for the fact that the “shock” happens at the branch level, and credit lines from the same branch might not be independent, we cluster the standard errors at the branch level in the estimations. We include branch fixed effects in some of our tests to account for any time-invariant branch-specific factors.

In equation (1), $D(Post)$ is an indicator that takes the value of 1 after the restructuring and the resultant shock (i.e., in the year 2006), and 0 otherwise. The indicator $D(Treated)$ takes the value of 1 for treated credit lines and 0 otherwise. The main parameter of interest

is the point estimate of the coefficient of the interaction term $D(Treated) \times D(Post)$, which captures the effect of the change in the allocation of decision-making authority on credit line terms. X denotes a set of control variables.

While our focus is on collateral, we also examine the interest rate in some of our tests. Hence, we construct an indicator *Collateral* that takes the value of 1 if a credit line is secured with collateral, and 0 otherwise. *Rate* is the interest rate, in percentage terms, charged by the bank. In Table 1, we present summary statistics of the outcomes of interest for the treated and control groups. Panel A shows mean and median values for *Collateral* and *Rate* before and after the change in decision-making authority, i.e., for year 2004 (Pre) and 2006 (Post). On average, the incidence of collateral, *Collateral*, is 40% in 2004 and 50% in 2006 for the treated credit lines. This indicates a modest increase in the use of collateral. For comparison, the incidence of collateral in the control group (Control 1) increases from 34% in 2004 to 58% in 2006 on average. Using the control group, we infer that the DD estimate of the effect on the incidence of collateral is negative.

Table 1. Summary statistics

The table presents summary statistics for the variables used in the analysis. Treated Group consists of all treated credit lines. Control Group consists of credit lines that are in the treatment range but are not at a treated branch (Control 1). *Collateral* is an indicator that takes the value of 1 if the credit line is secured with collateral, and 0 otherwise. *Rate* is the interest rate charged by the bank, expressed in percentage terms. Variable descriptions are provided in Appendix A.

Panel A					
		Pre		Post	
		Mean	Median	Mean	Median
Treated Group (N = 48)	Collateral	0.40	0.00	0.50	0.50
	Rate	6.38	5.66	6.96	6.42
Control Group (N = 220)	Collateral	0.34	0.00	0.58	1.00
	Rate	6.11	5.33	6.87	6.34

Panel B				
		Treated Group		Control Group
		Mean	Median	Mean
Sales		2.85	3.00	3.10
D(Sales 1)		0.06	0.00	0.04
D(Sales 2)		0.25	0.00	0.20
D(Sales 3)		0.48	0.00	0.41
D(Sales 4)		0.19	0.00	0.30
D(Sales 5)		0.02	0.00	0.05
D(Sales 6)		0.00	0.00	0.00
D(Sales 7)		0.00	0.00	0.00

	Treated Group		Control Group	
Rel Length	8.28	8.44	8.07	8.25
Multiple	0.89	1.00	0.92	1.00
Other Services	0.94	1.00	0.95	1.00
Borrower-Bank Distance	7.62	8.03	7.75	7.84
Borrower-Rivals Distance	7.90	8.05	8.04	7.97
D(Cluster)	0.67	1.00	0.70	1.00

Turning to the cost of debt, we note that the average interest rate for the treated group increases by 58 basis points (bps), from 6.38% in 2004 to 6.96% in 2006. The average interest rate for the credit lines in the control group also increases, from 6.11% in 2004 to 6.87% in 2006. Hence, the DD estimate of the effect on the rate charged by the bank is also negative.

Extant research suggests that various borrower characteristics affect contract terms. Hence, we include a set of control variables in our model. First, we control for borrower size using sales (Berger and Udell, 1995; Brick and Palia, 2007; Ono and Uesugi, 2009). Our bank does not provide actual sales figures but only sales bins ranging from 1 (smallest) to 7 (largest). Therefore, we create an indicator $D(\text{Sales } j)$ that takes the value of 1 if the sales of a borrower fall in the j -th sales bin, and 0 otherwise.⁶ We also include three characteristics of the lending relationship between a borrower and the bank, motivated by research that the nature of this relationship is an important determinant of collateral (Berger and Udell, 1995; Degryse and Van Cayseele, 2000; Elsas and Krahnen, 2000; Chakraborty and Hu, 2006; Brick and Palia, 2007; Ono and Uesugi, 2009; Berger et al., 2011; Bellucci et al., 2010; Bharath et al., 2011; Calcagnini et al., 2015). First, we consider the length of the relationship, *Rel Length*, measured as the natural logarithm of 1 plus the number of days since the borrower first started business with the bank. Second, we include the breadth of the relationship, which we capture using an indicator *Other Services*. It takes the value of 1 if a borrower uses other services, such as a checking account from the bank, and 0 otherwise. Last, we account for exclusivity of the relationship. The indicator *Multiple* takes the value of 1 if a borrower uses multiple banks for credit, and 0 if the borrower maintains an exclusive lending relationship with our bank.

Existing research also suggests that geographic and organizational distances capture factors relevant for loan contracting (Bellucci et al., 2019; Jimenez et al., 2006; Jimenez et al., 2009). As a result, we include the variable *Distance Borrower-Bank* measured as the natural logarithm of the metric distance between the borrower and the lending branch. We also include the distance between the borrower and the branches of other banks operating in

⁶The sales categories actually range from 1 to 9, but the credit lines in the sample used for the estimations are such that sales do not exceed category 7. Hence, we focus here on the relevant set of indicators.

the local credit market and competing with our bank. Specifically, we calculate the metric distance between a borrower and each branch of each bank present in the credit market, and take the natural logarithm of the median distance to construct *Distance Borrower-Rivals*. Lastly, we construct an indicator $D(Cluster)$ that takes the value of 1 if a borrowing firm is headquartered in an industrial cluster of economic activity and 0 otherwise.

Summary statistics for the control variables are provided in Panel B of Table 1. We present averages for both years for the treatment and control groups. We note that the borrowers with “treated” credit lines are similar in terms of size to those in the control group (Control 1). On average, “treated” borrowers have slightly longer lending relationships with the bank. The other two characteristics of their lending relationships are mostly comparable to those of the borrowers in the control group. Lastly, we note that, on average, borrowers in the treated group tend to be closer to the lending branch but also to branches of rival banks. These borrowers are slightly less likely to be part of industrial clusters. Thus, both groups seem fairly similar in observable characteristics.

3 Results

3.1 Baseline model

We begin our analysis by estimating a baseline specification of Equation (1) that includes only the indicators $D(Post)$ and $D(Treated)$, and their interaction. We then augment the model with control variables to ensure the estimated effect does not simply reflect factors related to the borrower or the nature of the bank-borrower interaction, and fixed effects. The standard errors are clustered at the branch level. As we have a relatively large number of fixed effects and to facilitate interpretation, we use OLS regressions for the main analysis, but we verify the robustness of our results to estimation methods that reflect the dichotomous nature of the dependent variable.

For our main analysis, the sample consists of credit lines within the treatment range granted by branches affected by the organizational restructuring (treated group) and credit lines within the treatment range but not from a treated branch (control group, or Control 1). The estimation results are presented in Table 2. The dependent variable in all columns is *Collateral*.

The results from estimating the baseline model are in column (1). Our analysis reveals that the increase in information availability resulting from the “transition” of the treated credit lines from the central to the local level is associated with a reduction in the use of collateral. The coefficient on the interaction term $D(Treated) \times D(Post)$ is negative and

statistically significant at 5% level. We estimate a reduction in collateral use of about 14.1%. This is economically important given the 40% unconditional incidence of collateral in the treatment group in the pre-treatment period.

In column (2), we augment the baseline model with a set of control variables and industry fixed effects. Consistently, we find that the coefficient of the interaction term is negative and statistically significant at 5% level. The magnitude of the estimated effect remains largely unchanged. Specifically, we estimate a 15% reduction in the incidence of collateral.

Table 2. Collateral and information: Baseline estimations

The table presents results from OLS models estimating the effect of changes in information availability on collateral. The sample consists of treated credit lines and credit lines in the control group Control 1 (credit lines in the treatment range from untreated branch). The dependent variable *Collateral* is an indicator that takes the value 1 if the credit line is secured by collateral and 0 otherwise. The table presents coefficient estimates followed by robust standard errors clustered at the branch level in parentheses. Variable descriptions are provided in Appendix A. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
D(Post)	0.245*** (0.037)	0.237*** (0.040)	0.245*** (0.038)	0.257*** (0.045)
D(Treated)	0.059 (0.045)	0.067 (0.059)		
D(Treated)×D(Post)	-0.141** (0.063)	-0.150** (0.065)	-0.141** (0.065)	-0.154** (0.066)
D(Sales 2)		0.031 (0.160)		0.145 (0.212)
D(Sales 3)		-0.018 (0.155)		0.121 (0.205)
D(Sales 4)		-0.096 (0.164)		0.081 (0.224)
D(Sales 5)		-0.211 (0.160)		-0.141 (0.226)
Rel Length		-0.021 (0.029)		-0.055 (0.036)
Other Services		-0.160 (0.155)		-0.274 (0.223)
Multiple		-0.106 (0.078)		-0.091 (0.074)
Borrower-Bank Distance		0.011 (0.026)		0.003 (0.031)
Borrower-Rivals Distance		-0.005		-0.008

	(1)	(2)	(3)	(4)
D(Cluster)		(0.026) 0.040		(0.030) -0.436**
Constant	0.336*** (0.039)	(0.056) 0.693* (0.397)	0.448*** (0.027)	(0.184) 1.046** (0.425)
Industry FE	No	Yes	No	Yes
Branch FE	No	No	Yes	Yes
Observations	536	536	536	536
Adjusted R-squared	0.047	0.112	0.080	0.153

In columns (3) and (4), we repeat the estimation of the models in columns (1) and (2) but add branch fixed effects to absorb all time-invariant branch characteristics. While the size-matched control group accounts for possible differences in collateral use related to the amount of extended credit, it does not account for branch-specific factors. We note that the estimated effects in columns (3) and (4) are consistent with our result that an increase in the overlap between decision-making authority and locus of information production is associated with a reduction in the use of collateral.

3.2 Underlying assumptions

A key premise of the DD analysis is the presence of parallel pre-trends. While our limited data prevent us from thoroughly evaluating the premise, we conduct a test that may be informative. Specifically, we estimate a version of the model outlined in equation (1) year by year. In this setting, we drop the indicator $D(Post)$ and its interaction with $D(Treated)$. We retain $D(Treated)$ and the controls. We have more than 20 industry indicators in the specification with industry fixed effects. This relatively large number of variables reduces power in the year-by-year analysis. Hence, we group industries into macro-sectors of economic activity and incorporate them in the model through sector fixed effects. The results of the estimations are reported in Table 3.

Table 3. Parallel trends

The table presents the results of year-by-year OLS model estimates of the effect of changes in information availability on collateral. Year 2004 (2006) is pre-restructuring (post-restructuring). The sample consists of treated credit lines and credit lines in the control group Control 1 (credit lines in the treatment range from untreated branch). The dependent variable *Collateral* is an indicator that takes the value 1 if the credit line is secured by collateral and 0 otherwise. The table presents coefficient estimates, with robust standard errors clustered at the branch level in parentheses. Controls are the same as in Table 2. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

	(1)	(2)
	Year 2004	Year 2006
D(Treated)	0.049 (0.051)	-0.093* (0.050)
Controls	Yes	Yes
Sector FE	Yes	Yes
Observations	268	268
Adjusted R-squared	0.020	0.006

Column (1) of the table shows the estimation results for the year 2004, while column (2) is for the year 2006. For the pre-treatment period, the coefficient on $D(Treated)$ is not statistically significant. This is consistent with the parallel trends assumption and suggests that the incidence of collateral is comparable across borrowers in the treatment and control groups pre-treatment. By contrast, the coefficient on $D(Treated)$ in column (2) is negative and significant at the 10% level. Thus, the use of collateral in treated credit lines is lower post-treatment. Thus, the cross-sectional estimation offers results consistent with the assumption underlying the DD approach.

3.3 Estimation method and additional control variables

Our main analysis is based on OLS estimations. We also estimate Logit models to take into account the dichotomous nature of the dependent variable *Collateral*. The estimation results are in column (1) of Table 4. We note that our findings are robust to this alternative estimation method.

The policy adopted by the bank for assigning credit lines to management at the local or central level relies on thresholds based on the credit line limit. The credit amount, however, can be determined simultaneously with other terms of the contract. Consequently, in our main analysis, we exclude contract terms such as amount from the set of controls. To ensure that our results do not mask credit-amount effects, we include *Amount*, which is the natural logarithm of the credit line limit. In column (2) of the table, we present the results of this analysis and show that our conclusions about collateral continue to hold.

Borrower risk affects both credit availability and contract terms. Machauer and Weber (1998) and Berger et al. (2011), among others, demonstrate that the use of ratings or credit scoring technology can be an important determinant of collateral. We explore this by leveraging information on internal ratings assigned by the bank to borrowers. However, we only have such information for a subset of borrowers. With this caveat in mind, we create

an indicator $D(Rated)$ that takes the value of 1 if a borrower has an internal rating, and 0 otherwise. We use the actual numerical score, coded as 0 if missing, to construct an integer variable $Rating$ that reflects the actual score. The estimation results of the models augmented with $D(Rated)$ and $Rating$ presented in columns (3) and (4) of the table demonstrate that our insights continue to hold.

Table 4. Additional controls and estimation methods

The table presents results from alternative specifications and models that estimate the effect of changes in information availability on collateral. The sample consists of all treated credit lines and the credit lines in the control group Control 1 (credit lines in the treatment range from untreated branch). The estimation method is Logit in column (1), while it is OLS in columns (2) to (4). The dependent variable *Collateral* is an indicator that takes the value 1 if the credit line is secured by collateral and 0 otherwise. The estimation in column (2) controls for credit line size (*Amount*). The estimations in columns (3) and (4) control for borrower risk by including 1) an indicator *D(Rated)* that takes a value of 1 for rated borrowers or 2) the numeric rating score *Rating* assigned by the bank. The table presents coefficient estimates followed by robust standard errors clustered at the branch level in parentheses. Variable descriptions are provided in Appendix A. Controls are the same as in Table 2. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
D(Post)	1.111*** (0.175)	0.203*** (0.040)	0.239*** (0.040)	0.239*** (0.040)
D(Treated)	0.337 (0.284)	0.064 (0.052)	0.064 (0.064)	0.062 (0.066)
D(Treated)×D(Post)	-0.749*** (0.288)	-0.097* (0.056)	-0.151** (0.064)	-0.152** (0.064)
Amount		0.403*** (0.051)		
D(Rated)			0.064 (0.043)	
Rating				0.015** (0.007)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	522	536	536	536
R-squared/Pseudo R-squared	0.114	0.248	0.168	0.172

3.4 Alternative control group

Given that the thresholds are based on credit line size, the control group in the main analysis (Control 1) is constructed so that untreated credit lines fall within the treatment size limits. Hence, the control group in our main analysis consists of credit lines originating from branches not subject to restructuring. It is possible that treated branches differ from untreated branches. While the use of branch fixed effects accounts for time-invariant branch characteristics, we also conduct a robustness test using an alternative control group that focuses only on the branches that undergo restructuring. Specifically, this control group (Control 2) includes all untreated credit lines originated at the treated branches and man-

aged locally. In other words, these are credit lines at the restructured branches that are below the treatment range. With this approach, all branch-specific factors are held constant. By design, the credit lines in this control group are smaller lines, and it is important to control for credit line size in the estimations. We include *Amount*, which is the natural logarithm of the credit line limit. Last, we note that the treatment group remains the same.

The results of the analysis with the alternative control group (Control 2) are in Table 5. Column (1) shows the results of the baseline specification that includes only the indicators $D(Post)$, $D(Treated)$, and their interaction. In column (2), we augment the baseline specification with all control variables and industry fixed effects, while in column (3), we also add branch fixed effects.

Table 5. Alternative control group

The table presents results from OLS models estimating the effect of changes in information availability on collateral. The sample consists of treated credit lines and an alternative control group Control 2 (untreated credit lines from the treated branches). Year 2004 (2006) is pre-restructuring (post-restructuring). The dependent variable *Collateral* is an indicator that takes the value 1 if the credit line is secured by collateral and 0 otherwise. The table presents coefficient estimates, with robust standard errors clustered at the branch level in parentheses. Variables descriptions are provided in Appendix A. Controls are the same as in Table 4. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	All	All	Year 2004	Year 2006	Year 2004	Year 2006
D(Post)	0.317*** (0.030)	0.319*** (0.022)	0.320*** (0.022)				
D(Treated)	0.260*** (0.024)	0.182** (0.039)	0.182** (0.037)	0.121* (0.042)	0.024 (0.061)	0.122* (0.040)	0.023 (0.062)
D(Treated)×D(Post)	-0.213* (0.071)	-0.215** (0.058)	-0.216** (0.058)				
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	No	No	Yes	No	No	Yes	Yes
Observations	1,958	1,958	1,958	979	979	979	979
Adjusted R-squared	0.118	0.212	0.213	0.160	0.117	0.161	0.117

Our analysis consistently shows that greater overlap between decision-making authority and information availability reduces the use of collateral. The coefficient on the interaction term $D(Treated) \times D(Post)$ is negative and statistically significant in columns (1) to (3). Depending on the specification, the estimated reduction in use of collateral is about 21% to 22%. Thus, we infer that our insights are not driven by the control group we adopt.

We also repeat the year-by-year analysis using the alternative control group. The estimation results for the year 2004 are in columns (4) and (6), while those for the year 2006 are in columns (5) and (7). We note from columns (4) and (6) that pre-restructuring, the treated credit lines are more likely to be secured with collateral. The coefficient estimate of $D(Treated)$ is positive and significant at the 10% level. By contrast, the estimates in columns (5) and (7), which are based on post-restructuring data, are not statistically significant. The magnitude of the coefficient also drops substantially. This is consistent with the notion that once credit lines are transferred from the central to the local level, their characteristics converge to those of other locally managed contracts.

4 Underlying mechanisms and moderating effects

4.1 Divergence of information within the bank organization

Our analysis demonstrates that an increase in the overlap between decision-making authority and information production is robustly associated with reduced use of collateral. We now conduct tests to show that this relationship reflects the loan officers' incentives to use more soft information once they gain formal decision-making authority about loan approval. More precisely, we explore factors that can moderate the estimated effect and point to the underlying mechanisms.

As information frictions between lender and borrower are central to our argument, we first consider the role of borrowers' informational opacity and information available to the loan officer. If information is an important driver of the estimated effect of transferring contracts from the central to the local level, the effect should be stronger when the information sets of loan officers operating at the local and central levels are more divergent. This is more likely to occur when soft information exists or when local loan officers have tacit local knowledge unavailable to central-level officers. The underlying premise is that officers at the lending branch are the repository of soft information and tacit local knowledge. By contrast, the effect should be weaker when soft information or local knowledge is limited.

To implement our tests, we first capture variation in the amount of information using the length of the lending relationship between the borrower and the bank. Specifically, we construct an indicator $D(Short\ Rel\ Length)$ that takes the value 1 if the relationship length falls in the bottom tercile of the sample distribution, and 0 otherwise. For such borrowers, the bank has not yet accumulated much soft information, and the information sets of officers at the local and central levels are similar. Hence, authority allocation should be less relevant, and the estimated effect should be smaller. To test the conjecture, we augment the model

with a triple interaction $D(Treated) \times D(Post) \times D(Short\ Rel\ Length)$, along with all double interaction terms needed to ensure a comprehensive specification. The results are presented in column (1) of Table 6.

In these models, the coefficient on the double interaction $D(Treated) \times D(Post)$ reflects the estimated effect for borrowers with a longer lending relationship. The point estimate in column (1) is negative and significant at 5% level. Thus, for borrowers with longer lending relationships, the reallocation of authority reduces reliance on collateral. The coefficient of the triple interaction is positive. To compute the effect for borrowers with a shorter lending relationship, we sum the coefficients of $D(Treated) \times D(Post) \times D(Short\ Rel\ Length)$ and $D(Treated) \times D(Post)$. The linear combination is not statistically significant, suggesting that, for borrowers with short lending relationships, the transfer of authority does not have a significant effect. Thus, we infer that the effect is driven by cases with potential divergence in the information sets of local and central levels due to soft information accumulated over the course of the longer lending relationship.

Divergence in information sets of different levels can also stem from the physical proximity between borrowers and local bank branches. Local loan officers can have relevant tacit knowledge when they are placed close to borrowers, and this knowledge can be similarly difficult to transfer to the central level. To examine this notion, we construct an indicator $D(Long\ Distance)$ that takes the value 1 if the distance between borrower and lending branch is in the top tercile of the sample distribution, and 0 otherwise. We then augment the model with the triple interaction $D(Treated) \times D(Post) \times D(Long\ Distance)$ and all relevant double interactions. In this setting, the coefficient on the double interaction $D(Treated) \times D(Post)$ captures the effect for borrowers located close to the lending branch, where proximity could create differences in the information sets of local and central agents due to tacit local knowledge.

The estimated coefficient on $D(Treated) \times D(Post)$ is negative and significant at the 1% level in column (2). This suggests that a change in decision-making authority reduces the use of collateral when the borrower is located near a local bank branch. The effect loses significance for more distant borrowers. The sum of the coefficients of $D(Treated) \times D(Post)$ and $D(Treated) \times D(Post) \times D(Long\ Distance)$ is not different from 0, as indicated by the p-value for the linear combination of the two terms. Thus, we find no effect in cases where the information available to the local and central levels is comparable, because the borrower is located far from the local loan officer and the lender's tacit local knowledge may not be applicable.

Table 6. Information availability and authority allocation

The table presents results of OLS models that estimate the effect of the change in information availability on collateral, conditional on 1) available information and 2) functional distance. The sample consists of all treated credit lines and credit lines in the control group Control 1 (credit lines in the treatment range from untreated branch). The dependent variable *Collateral* is an indicator that takes value of 1 if the credit line is secured with collateral, and 0 otherwise. *D(Short Rel Length)* is an indicator that takes value of 1 if the lending relationship between borrower and bank is in the bottom tercile of the sample distribution, and 0 otherwise. *D(Long Distance)* is an indicator that takes value of 1 if the distance between borrower and bank branch is in the top tercile of the sample distribution, and 0 otherwise. *D(Short Functional Distance)* is an indicator that takes the value 1 if the distance between a branch and the bank headquarters is in the bottom tercile of the sample distribution, and 0 otherwise. The table presents coefficient estimates, with robust standard errors clustered at the branch level in parentheses. Variable descriptions are provided in Appendix A. Controls are the same as in Table 2. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

		(1)	(2)	(3)
D(Post)		0.255*** (0.051)	0.228*** (0.049)	0.236*** (0.051)
D(Treated)		0.084 (0.069)	0.117 (0.115)	0.075 (0.048)
D(Treated)×D(Post)	(A)	-0.171** (0.076)	-0.176*** (0.062)	-0.107 (0.074)
D(Treated)×D(Post)×D(Short Rel Length)	(B)	0.074 (0.139)		
D(Post)×D(Short Rel Length)		-0.057 (0.080)		
D(Treated)×D(Short Rel Length)		-0.053 (0.079)		
D(Short Rel Length)		0.081 (0.058)		
D(Treated)×D(Post)×D(Long Distance)	(C)		0.075 (0.074)	
D(Post)×D(Long Distance)			0.024 (0.072)	
D(Treated)×D(Long Distance)			-0.146 (0.279)	
D(Long Distance)			0.023 (0.073)	
D(Treated)×D(Post)×D(Short Funct. Dist.)	(D)			-0.154* (0.090)
D(Post)×D(Short Functional Distance)				0.010 (0.072)
D(Treated)×D(Short Functional Distance)				0.012 (0.067)
D(Short Functional Distance)				0.126** (0.047)

	(1)	(2)	(3)
Rel Length		-0.019 (0.028)	-0.022 (0.031)
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	536	536	536
Adjusted R-squared	0.109	0.109	0.118
Linear combination (A) + (B)	-0.097		
<i>p-value</i>	<i>0.527</i>		
Linear combination (A) + (C)		-0.102	
<i>p-value</i>		<i>0.206</i>	
Linear combination (A) + (D)			-0.261
<i>p-value</i>			<i>0.001</i>

Overall, our analysis reveals that the allocation of decision-making authority has a stronger effect when the potential for divergence between the information sets of the local and central levels is greater. Moreover, given that the effect is realized when soft information is likely to have been accumulated by the local loan officer (e.g., over the course of the lending relationship), our results are also consistent with the idea that delegation of authority to local loan officers improves their incentives to fully utilize available information rather than to produce previously unavailable information.

4.2 Real and formal authority

The restructuring initiative implemented by the bank shifted formal authority to local loan officers. However, the effect of the shift likely depends on the pre-restructuring real authority enjoyed by the local loan officers. Extant research suggests that principals can have formal but not real authority if their decisions depend on agent-generated information that is difficult to evaluate (e.g., Aghion and Tirole, 1997). Thus, even without formal authority, local loan officers could influence the outcome of the contracting process if senior managers' decisions are based on the input and credit recommendations provided by the local loan officers. In this case, the allocation of formal authority might not have a significant effect as the contract terms already reflect the real authority of the local loan officers. By contrast, if they do not have significant real authority, granting formal authority might be associated with a differential outcome.

To examine whether allocation of formal authority interacts with the real authority of the loan officers at the local branches, we test for heterogeneous effects based on functional

distance, defined as the distance between the lending branch and the bank headquarters. We construct an indicator $D(\text{Short Functional Distance})$ that takes value of 1 if the distance between a branch and the headquarters is in the bottom tercile of the sample distribution, and 0 otherwise. We then add the triple interaction $D(\text{Treated}) \times D(\text{Post}) \times D(\text{Short Functional Distance})$ to the baseline model. In the augmented model, the coefficient on $D(\text{Treated}) \times D(\text{Post})$ captures the effect of the allocation of formal decision-making authority to branches operating far from headquarters. The sum of the coefficients of $D(\text{Treated}) \times D(\text{Post})$ and $D(\text{Treated}) \times D(\text{Post}) \times D(\text{Short Functional Distance})$ shows the effect for branches closer to the headquarters.

Column (3) of Table 6 shows that the coefficient of $D(\text{Treated}) \times D(\text{Post})$ is not significant. The sum of the coefficients of $D(\text{Treated}) \times D(\text{Post})$ and $D(\text{Treated}) \times D(\text{Post}) \times D(\text{Short Functional Distance})$ is negative and significant at 1% level. We infer that the allocation of formal authority is more likely to reduce the use of collateral when functional distance is shorter, i.e., when the lending branch is located close to the headquarters. In this case, the real authority of the loan officers prior to the allocation of formal authority is more limited. By contrast, when local loan officers are likely to have more real authority, as they are far from headquarters, the allocation of formal authority is not associated with a differential use of collateral. Thus, we conclude that when real authority is limited, the allocation of formal authority substantially affects the use of collateral.

5 Interest rate

Our analysis focuses on collateral as a contractual mechanism intended to mitigate information frictions between lenders and borrowers. As the transfer of decision-making authority from central to local level increases the overlap between the production of information and the ability to act on it, we consistently find that this increase is associated with a reduction in the use of collateral. While we focus on this contract term, other contract terms may also be affected because they are complements or substitutes for collateral.

We examine the cost of credit reflected in the interest rate charged by the bank. Specifically, we estimate Equation (1) using *Rate* as the dependent variable. We replicate the main analysis based on Table 2, and the examination of heterogeneous effects based on Table 6, and report the estimation results in Table 7.

By and large, we do not find a significant effect of allocating authority to local loan officers on interest rates, on average. In columns (1) through (5) of the table, the coefficient of the interaction term $D(\text{Treated}) \times D(\text{Post})$ is not statistically significant. Thus, it seems that local loan officers do not substitute collateral for higher rates. We also cannot detect

Table 7. Interest rate

The table presents the results of OLS models estimating the effect of changes in information availability on the interest rate. With the exception of column (5), the sample consists of treated credit lines and credit lines in the control group Control 1 (credit lines in the treatment range from untreated branch). In column (5), the sample consists of treated credit lines and an alternative control group, Control 2 (untreated credit lines that are from the treated branches). The dependent variable *Rate* is the interest rate charged by the bank, expressed as a percentage. The table presents coefficient estimates, with robust standard errors clustered at the branch level in parentheses. Variable descriptions are provided in Appendix A. Controls as in Table 2. *, **, and *** indicate statistical significance at 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D(Post)	0.754*** (0.116)	0.736*** (0.110)	0.754*** (0.121)	0.762*** (0.124)	0.887*** (0.080)	0.640*** (0.095)	0.886*** (0.125)	0.776*** (0.166)
D(Treated)	0.268 (0.393)	0.422 (0.400)			0.277 (0.327)	-0.269 (0.169)	0.141 (0.383)	0.654 (0.551)
D(Treated)×D(Post) (A)	-0.176 (0.446)	-0.210 (0.447)	-0.176 (0.464)	-0.227 (0.464)	-0.274 (0.439)	0.446 (0.414)	-0.298 (0.310)	-0.75*** (0.206)
D(T)×D(P)×D(Short R Length) (B)						-1.529 (0.935)		
D(Post)×D(Short Rel Length)						0.091 (0.276)		
D(T)×D(Short Rel Length)						1.686** (0.665)		
D(Short Rel Length)						-0.121 (0.306)		
D(T)×D(P)×D(Long Distance) (C)							0.292 (0.533)	
D(Post)×D(Long Distance)							-0.445 (0.309)	
D(Treated)×D(Long Distance)							0.622 (0.738)	
D(Long Distance)							0.622* (0.369)	
D(T)×D(P)×D(Short Fun Dist) (D)								1.970*** (0.246)
D(Post)×D(Short Funct Dist)								-0.116 (0.218)
D(T)×D(Short Functional Dist)								-0.784 (0.605)
D(Short Functional Dist)								0.271 (0.321)
Controls	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Branch FE	No	No	Yes	Yes	No	No	No	No
Observations	536	536	536	536	1,958	536	536	536
Adjusted R-squared	0.033	0.101	0.177	0.309	0.057	0.108	0.103	0.107
Linear combination (A) + (B) <i>p-value</i>						-1.083 0.251		
Linear combination (A) + (C) <i>p-value</i>							-0.007 0.993	
Linear combination (A) + (D) <i>p-value</i>								1.221 0.001

significant interactions between the allocation of decision-making authority and measures of potential divergence in the information sets at the local and central levels.

However, we find that the allocation of formal authority might differentially affect rates based on the real authority of the loan officers. In column (8), the coefficient of the interaction term $D(Treated) \times D(Post)$ is negative and significant at the 1% level. We infer that when the lending branch is far from the headquarters and the local loan officers have real authority over the lending process, the allocation of formal authority is associated with a reduction in interest rates. The linear combination of $D(Treated) \times D(Post)$ and $D(Treated) \times D(Post) \times D(Short\ Functional\ Distance)$ is positive and significant at the 1% level. To interpret, when real authority is limited, the allocation of formal authority is associated with an increase in interest rates. Thus, it seems that in this case, the reduction in the use of collateral is accompanied by an increase in the price of credit.

6 Conclusion

This paper explores how changes in the allocation of decision-making authority and in the information available to the lender affect contract terms in small-business lending. Our analysis exploits largely exogenous variation resulting from the restructuring initiative implemented by a medium-sized regional Italian bank. As part of the restructuring, the bank modified its assignment policy, which determines whether a credit line contract is managed locally at the originating branch by the local loan officer or transferred to the bank's headquarters, where senior bank managers make decisions. Specifically, the thresholds at some of the branches were doubled. Consequently, some credit lines previously managed centrally at the bank headquarters were returned to the local bank branches. This increased the information available to the loan officer responsible for managing the credit line. In other words, it increased the overlap between information production and the ability to act on it.

Using a DD estimation strategy, we show that an exogenous shock that increases the authority of the loan officer to produce information about the borrower is associated with a reduction in the use of collateral. By contrast, we do not observe meaningful effects on the price of credit reflected in the interest rate charged by the bank. Exploring the underlying drivers of the estimated effect on collateral requirements, we show that the effect is more pronounced when relevant tacit local knowledge and soft information are more likely to exist. On the contrary, allocating formal decision-making authority to local loan officers has a limited effect when the officers already exert real authority over the lending process. Thus, we highlight the complex interplay between decision-making authority and information production in the context of bank lending to SMEs.

References

- Agarwal, S., and R. Hauswald, 2010, Authority and information. *AFA 2011 Meetings Paper*. Available at SSRN: <https://ssrn.com/abstract=1566062>
- Aghion, P., and J. Tirole, 1997, Formal and real authority in organizations, *Journal of Political Economy* 105, 1-29.
- Albareto, G., M. Benvenuti, S. Mocetti, M. Pagnini, and P. Rossi, 2011, The organization of lending and the use of credit scoring techniques in Italian banks, *Journal of Financial Transformation, Applied Finance* #32, 133-142.
- Alessandrini, P., A. Presbitero, and A. Zazzaro, 2009, Banks, distances and firms' financing constraints, *Review of Finance*, 13(2), 261-307.
- Alonso, R., D. Dessein, and N. Matouschek, 2008, When does coordination require centralization?, *American Economic Review* 98, 145-179.
- Bellucci A., A. Borisov, and A. Zazzaro, 2010, Does gender matter in ban-firm relationships? Evidence from Small Business Lending, *Journal of Banking and Finance* 34, 2968-2984.
- Bellucci, A., A. Borisov, G. Giombini, and A. Zazzaro, 2019, Collateralization and distance, *Journal of Banking and Finance* 100, 205-217.
- Berger, A., and G. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 68, 351-381.
- Berger, A., M. Espinosa-Vega, W. Frame, and N. Miller, 2011, Why do borrowers pledge collateral? New empirical evidence on the role of asymmetric information, *Journal of Financial Intermediation* 20, 55-70.
- Besanko, D., and A. Thakor, 1987, Collateral and rationing: sorting equilibria in monopolistic and competitive credit markets, *International Economic Review* 28, 671-689.
- Bharath S., S. Dahiya, A. Saunders, and A. Srinivasan, 2011, Lending relationships and loan contract terms, *Review of Financial Studies* 24, 1141-1203.
- Boot, A., A. Thakor, and G. Udell, 1991, Secured lending and default risk: equilibrium analysis, policy implications and empirical results, *Economic Journal* 191, 458-472.
- Boot, A., and A. Thakor, 1994, Moral hazard and secured lending in an infinitely repeated credit market game, *International Economic Review* 35, 899-920.
- Bouwens, J., and T. de Kok, 2021, Can centralization improve the use of soft information? Evidence from a field study and lab experiment, *Working paper*. Available at SSRN: <https://ssrn.com/abstract=2823253>
- Bouwens, J., and P. Kroos, 2019. The effect of delegation of decision rights and control: The case of lending decisions for small firms, *Management Accounting Research* 43, 29-44.

Brick I., and D. Palia, 2007, Evidence of jointness in the terms of relationship lending, *Journal of Financial Intermediation* 16, 452-476.

Calcagnini, G., R. Cole, G. Giombini, and G. Grandicelli, 2018, Hierarchy of bank loan approval and loan performance, *Economia Politica* 35, 935-954.

Calcagnini G., G. Giombini, and E. Lenti, 2015, Gender differences in bank loan access. An empirical analysis, *Italian Economic Journal* 1, 193-217.

Canales, R., and R. Nanda, 2012, A darker side to decentralized banks: Market power and credit rationing in SME lending, *Journal of Financial Economics* 105, 353-366.

Chakraborty A., and C. X. Hu, 2006, Lending relationships in line-of-credit and non-line-of-credit loans: Evidence from collateral use in small business, *Journal of Financial Intermediation* 15, 86-107.

Chan, Y., and G. Kanatas, 1985, Asymmetric valuations and the role of collateral in loan agreements, *Journal of Money, Credit, and Banking* 17, 84-95.

Chan, Y., and A. Thakor, 1987, Collateral and competitive equilibria with moral hazard and private information, *Journal of Finance* 42, 345-363.

Degryse, H., and P. Van Cayseele, 2000, Relationship lending within a bank-based system: evidence from European small business data, *Journal of Financial Intermediation* 9, 90-109.

Dessein, W., 2002. Authority and communication in organizations. *The Review of Economic Studies* 69, 811-838.

Dell’Ariccia, G., and R. Marquez, 2006, Lending booms and lending standards, *Journal of Finance* 61, 2511-2546.

Elsas, R., and J. Krahnen, 2000, Collateral, default risk, and relationship lending: an empirical study on financial contracting, *CEPR Discussion Paper* #2540.

Filomeni, S., G. Udell, and A. Zazzaro, 2021, Hardening soft information: does organizational distance matter? *The European Journal of Finance* 27, 897-927.

Hattori, M., K. Shintani, and H. Uchida, 2015, The repository of soft information within bank organizations, *Journal of Money, Credit and Banking* 47, 737-770.

Jimenez, G., V. Salas, and J. Saurina, 2006, Determinants of collateral, *Journal of Financial Economics* 81, 255-281.

Jimenez, G., V. Salas, and J. Saurina, 2009, Organizational distance and use of collateral for business loans, *Journal of Banking and Finance* 33, 234-243.

Liberti, J. M., 2018, Initiative, incentives, and soft information, *Management Science* 64, 3714-3734.

Liberti, J. M., and A. Mian, 2009, Estimating the effect of hierarchies on information use, *Review of Financial Studies* 22, 4057-4090.

Liberti, J. M., and M. A. Petersen, 2019, Information: hard and soft, *Review of Corporate*

Finance Studies 8, 1-41.

Machauer, A., and M. Weber, 1998, Bank behavior based on internal credit ratings of borrowers, *Journal of Banking and Finance* 22, 1355-1383.

Mocetti, S., M. Pagnini, and E. Sette, 2017, Information technology and banking organization, *Journal of Financial Services Research* 51, 313-338.

Ono A., and I. Uesugi, 2009, Role of collateral and personal guarantees in relationship lending: Evidence from Japan's SME loan market, *Journal of Money, Credit and Banking* 41, 935-960.

Petersen, M., and R. Rajan, 1994, The benefits of lending relationships: evidence from small business data, *Journal of Finance* 47, 1367-1400.

Porzio, C., G. Sampagnaro, and V. Verdoliva, 2020, Lost in translation: The determinants and the effect of soft information dispersion in bank lending, *Global Finance Journal* 43, 100418.

Qian, Q. J., P. Strahan, and Z. Yang, 2015, The impact of incentives and communication costs on information production and use: Evidence from bank lending, *Journal of Finance* 70, 1457-1493.

Skrastins, J., and V. Vig, 2019, How organizational hierarchy affects information production, *Review of Financial Studies* 32, 564-604.

Stein, J., 2002, Information production and capital allocation: Decentralized versus hierarchical firms, *Journal of Finance* 57, 1891-1921.

A List of variables

Variable	Definition
Collateral	Indicator variable that takes the value 1 if the credit line is secured with collateral, and 0 otherwise.
Rate	Continuous variable that measures the interest rate on the credit line, expressed in percentage terms.
D(Post)	Indicator variable that takes the value 1 in the post-restructuring period (i.e., year 2006), and 0 otherwise.
D(Treated)	Indicator variable that takes the value 1 for treated credit lines, and 0 otherwise. A credit line is treated if 1) it is originated at a treated branch (i.e., a branch subject to restructuring and increased thresholds for credit line management) and 2) the amount of the credit line falls within the range of thresholds subject to restructuring (i.e., above the old (pre-2005) threshold and below the new (post-2005) threshold).
D(Sales j)	Indicator variable that takes the value 1 if the sales of a borrower fall in the j-th sales bin (j ranges from 1 to 7), and 0 otherwise. The sales bins are as follows: 1 for sales below €0.25M, 2 for sales between €0.25M and €0.5M, 3 for sales between €0.5M and €1.5M, 4 for sales between €1.5M and €5M, 5 for sales between €5M and €25M, 6 for sales between €25M and €50M, and 7 for sales that exceed €50M.
D(Short Rel Length)	Indicator variable that takes the value 1 if the lending relationship between the borrower and the bank is in the bottom tercile of the sample distribution, and 0 otherwise.
D(Long Distance)	Indicator variable that takes the value 1 if the distance between the borrower and the lending branch is in the top tercile of the sample distribution, and 0 otherwise.
D(Short Functional Distance)	Indicator variable that takes the value 1 if the distance between the lending branch, where the credit line originates, and the headquarters of the bank is in the bottom tercile of the sample distribution, and 0 otherwise.
D(Cluster)	Indicator variable that takes the value 1 if the borrower is headquartered in an industrial cluster, and 0 otherwise.
D(Rated)	Indicator variable that takes the value 1 if a borrower has internal rating by the bank, and 0 otherwise.
Rating	Integer variable that reflects the numerical rating score assigned to a borrower by the bank.
Rel Length	Natural logarithm of 1 + the length of the lending relationship (in days) between the bank and a borrower.
Multiple	Indicator variable that takes the value 1 if the borrower has lending relationship with multiple banks, and 0 if the borrower has an exclusive lending relationship with the bank.

Variable	Definition
Other Services	Indicator variable that takes the value 1 if the bank provides other services in addition to the credit line (e.g., checking account) to the borrower, and 0 otherwise.
Distance Borrower-Bank	Natural logarithm of the metric distance between the borrower and the branch where the credit line originates.
Distance Borrower-Rivals	Natural logarithm of the median of the metric distances between the borrower and the branches of other banks (i.e., rivals) operating in the local credit market.
Amount	Natural logarithm of the limit on the credit line, expressed in thousands.

B Background and External Validity

In this appendix, we provide additional information about bank organization, economic structure of the two provinces represented in our dataset, and characteristics of the sample to support the generalizability and external validity of our inferences. First, we argue that the results we obtain from the Italian context are relevant to other economies. In the last year of our sample period, the Italian banking sector consisted of more than 750 banks and was the fourth largest in Europe in terms of assets. The profitability and efficiency of Italian banks was generally comparable to that of French and German banks (Drummond et al., 2007). Moreover, similar to other countries, the Italian banking sector experienced a trend of consolidation over time (Papi et al., 2017). Second, our bank has characteristics and strategy similar to many banking institutions in Italy that exhibit significant regional presence but limited operations at national or international level. At the start of the period, the bank operated branches in 16 provinces and had substantial regional presence and local reach. Over time, the bank grew at the regional and local level through a strategy of acquisitions of small community banks. Third, the characteristics of the contracts in the sample we examine are broadly consistent with aggregate statistics for the Italian banking system. For instance, the share of loans secured with collateral in Italy in 2004 was about 39% (Bellucci et al., 2019). In our study, the share is 40% for the credit lines in the treated group and 34% for the credit lines in the control group (Control 1). Moreover, the credit lines are granted to borrowers operating in various sectors of economic activity and organizational forms. Specifically, we have borrowers from more than 20 sectors (the bank distinguishes sectors of economic activity using the 2-digit level classification of the Italian National Institute of Statistics) and the represented organizational forms are Sole Proprietorships, Partnerships, Corporations, and Cooperatives. Last, the provinces we study are representative of the Italian economic structure and conditions at the time according to macroeconomic indicators and industrial composition (Bellucci et al., 2010). For instance, using average GDP per capita in Italy in 2004 as a basis of 100, one of the provinces has slightly lower GDP per capita (96) and the other slightly higher (109). In both provinces manufacturing accounts for about 30% of economic activity in terms of value added, services add 68%, and agriculture is 2%. The figures for the overall Italian economy are 27%, 71%, and 2%, respectively.

Status seeking and positional consumption under mobility restrictions during COVID-19*

Vincenzo Lombardo[†] Angela Pepe

Abstract

This paper investigates the role of social visibility in consumption behavior by studying how mobility restrictions during the COVID-19 pandemic affected expenditures on visible and positional goods. Using detailed microdata from the Italian Household Budget Survey (HBS) combined with municipality-level excess mortality data, we document that visible consumption declined significantly during periods of restricted mobility. The results highlight the behavioral nature of conspicuous consumption and provide novel evidence on the link between visibility, social status, and consumption networks.

Keywords: COVID-19 pandemics, positional consumption, preference for status, peer effects

JEL: D12, D14, D85, E21, Z10, Z13

*Vincenzo Lombardo greatly acknowledges research funds received by the European Union – Next Generation EU, under the Age-It project - Ageing well in an ageing society (AGE-IT PE00000015 - CUP H73C22000900006) “FENICE - Fragile Elderly and eNerhy in Italian Consumption Empirics”. Authors also greatly thank the Italian National Institute of Statistics (ISTAT) for providing access to the microdata files of "The household budget survey (HBS)" (Codice Ist-02396 - Protocol ISTAT-DISAE ACP/24/2023).

[†]Department of Business and Economics (DISAE), University of Naples Parthenope. E-mail: vincentzo.lombardo@uniparthenope.it.

1 Introduction

Understanding why people consume what they consume is a question that runs through the core of economic theory and public policy. Since the pioneering work of Thorstein Veblen, economists have recognized that consumption goods do not merely satisfy private wants: they also convey information about status, identity, and social rank. Veblen's phraseology crystallized a basic observation that many subsequent authors have formalized: a substantial portion of consumption has a signalling dimension, where certain goods derive part of their value from being observed. When social interactions and visibility change, the incentives that underlie observable consumption must also change. Yet despite the extensive theoretical and empirical literature on status-related consumption, relatively little is known about how demand for visible goods responds when visibility itself is suddenly and substantially reduced.

The COVID-19 pandemic created precisely such a shock. Lockdowns, social distancing, and local outbreaks disrupted ordinary patterns of mobility and interaction, and thus abruptly curtailed individuals' opportunities to see and be seen. In Italy, 2020 witnessed large heterogeneity in the local intensity of the pandemic, both across municipalities and over months, offering a natural setting to examine whether reductions in visibility materially altered the composition of household spending. The intuition is straightforward: if part of the utility from purchasing certain goods depends on the chance that others will observe them, then limiting encounters should reduce the marginal value of such purchases and thus lower demand. If instead the appeal of those goods is intrinsic (independent of being observed), changes in interaction opportunities should have limited impact beyond income or availability effects.

This paper uses rich microdata from the Italian Household Budget Survey (*Indagine sulle Spese delle Famiglie*, HBS) for 2019–2020, combined with municipal excess mortality statistics from ISTAT and mobility indicators from Facebook's Movement Range program, to document and explain how the pandemic affected spending on goods that are plausibly visible in ordinary social interactions. The HBS is the official source for household consumption in Italy and provides monthly diaries with remarkably fine expenditure disaggregation (six-digit categories covering nearly 500 items), geographic identifiers at the municipality level, and household characteristic variables that allow for careful control of demographic and resource differences. We classify visible goods following the approach used in the literature: categories such as clothing and jewellery, personal care services, vehicles and vehicle-related goods, alcohol and

tobacco, and certain forms of entertainment and hospitality are characterised by high external observability and therefore are natural candidates for positional or conspicuous consumption.

To measure pandemic intensity, we rely on municipality-month excess mortality, defined as the difference between deaths in a given 2020 month and the corresponding average in 2015–2019. Excess mortality captures both the direct health burden of COVID-19 and, indirectly, the local behavioural and policy responses that constrained mobility and interaction. Facebook Movement Range data are used as an auxiliary measure to validate that months and places with higher excess mortality were also those where mobility declined most. While Facebook coverage is incomplete and thus not used as the primary explanatory variable, it confirms the key identifying assumption: excess mortality is strongly associated with reductions in movements and therefore with the practical erosion of opportunities to observe others' consumption.

Our empirical approach combines several complementary strategies to explore three related questions: First, did expenditures on visible goods decline more in places and months with more severe pandemic outbreaks? Second, do household expenditures on visible goods reflect local consumption norms, consistent with positional consumption? Third, did the pandemic moderate the strength of local social effects—so that in places where signalling was previously more important, the collapse of visibility induced larger shifts in spending? We address these questions by exploiting the geographic and temporal variation in the HBS combined with municipality characteristics and multiple robustness checks. The repeated cross-sectional design of the HBS prevents individual-level panel identification, but the large sample, monthly frequency, and fine spatial granularity provide rich quasi-experimental variation that is well-suited to the research tasks at hand.

The results are consistent and economically meaningful. Expenditure on visible goods fell significantly more in municipalities experiencing larger excess mortality, even after controlling for household total expenditure and fixed effects that absorb unobserved municipality-level factors and aggregate month-specific shocks. The timing of the decline closely tracks the months of strictest restrictions and largest mobility reductions. Importantly, categories that are not typically classified as visible display much smaller or negligible declines once the same controls are in place, suggesting that the effects are not driven solely by aggregate income shocks or by the closure of specific supply channels. We also document that households' spending on visible

goods is positively associated with the average visible consumption of other households in the same municipality, corroborating the notion that visible goods operate, in part, as positional goods tied to local reference groups. Finally, the reduction in visible consumption was most pronounced in municipalities where pre-pandemic visible consumption was higher, consistent with the idea that the social value of consumption was suppressed where signalling incentives were strongest.

The theoretical motivation for this empirical program is drawn from two complementary strands of thought in the literature on positional consumption. One strand emphasizes the role of consumption networks and social comparisons: individuals' utility depends on their consumption relative to that of peers, co-workers, and neighbours, and such interdependence can shape labour supply, saving, and aggregate outcomes. The other strand emphasizes deeper preference parameters—evolutionary, psychological, or culturally embedded inclinations toward status-seeking—that make status signals intrinsically valuable. While the former highlights the contextual nature of consumption (it depends on the behaviour of those around the individual), the latter underscores the more enduring dispositional motives that can sustain status-seeking even in the absence of immediate observability. Both perspectives imply that an exogenous, sudden reduction in opportunities for social observation ought to alter the demand for goods whose value is partly social. The pattern we document, sharp declines in visible goods concurrent with mobility reductions and larger effects where local signalling is typically more important, speaks directly to this mechanism.

Related literature. This paper intersects three literatures. First, it speaks to the large and growing empirical literature on status, relative consumption, and reference-group effects. Early empirical contributions exposed how consumption and well-being react to relative income and local consumption norms (Luttmer, 2005; Charles et al., 2009; Heffetz, 2011). Several studies have leveraged creative identification strategies to isolate social externalities: for instance, partner and co-worker reference groups have been used to address reflection and sorting concerns (De Giorgi et al., 2020), and natural experiments have documented signalling motives in contexts from educational choices to housing. Our work adds to this strand by focusing explicitly on how opportunities for visibility affect the positional dimension of consumption, using high-frequency microdata to capture both the timing and locality of adjustments.

Second, the paper contributes to the literature on the economic impacts of the COVID-19 pandemic. A broad set of studies has documented dramatic changes in

consumption, savings, labour supply, and expectations during 2020 (Baker et al., 2020; Barro et al., 2020; Christelis et al., 2020; Bachmann et al., 2024). Much of that literature has concentrated on aggregate spending patterns, liquidity constraints, or sectoral heterogeneity. By contrast, our analysis examines the micro-level reallocation within households' consumption baskets, with a focus on visibility and signaling motives; this perspective complements sectoral and aggregate accounts by showing how compositional shifts reflect responses to social interaction constraints.

Third, the paper relates to research that uses mobility and local epidemiological variation to study behavioral responses. Mobility measures, including those provided by tech platforms, have been used as crucial inputs for understanding compliance, economic activity, and the spread of the virus. We use such measures primarily as validation and robustness checks; our main identification relies on excess mortality because it more directly captures the local intensity of the shock and is available for the entire set of municipalities covered by the HBS.

The paper is structured as follows. After this introduction and literature discussion, we describe the HBS and our classification of visible goods, and we document descriptive correlations between excess mortality, mobility, and local consumption patterns. We then present the empirical strategy and a battery of robustness checks, including alternative definitions of visibility, placebo analyses, and checks for compositional change. The core empirical results show substantial and robust declines in visible consumption associated with pandemic intensity, evidence of positional effects at the municipality level, and heterogeneous responses that are concentrated where signalling motives were strongest before the pandemic. We conclude by discussing implications for theories of status-driven consumption and for policy: changes in social interaction patterns – whether temporary, as in the pandemic, or more persistent owing to technological and demographic shifts – may have far-reaching consequences for the composition of demand, savings, welfare, and inequality.

2 Data

Our main source of data is the the National Institute of Statistics (ISTAT) from which we derive information on the Italian households' expenditure from the Household Budget Survey (HBS) and on the excess mortality during the COVID-19 pandemics. We also use data on individuals' mobility from Facebook Data for Good as a robustness and validation of our main measure of the mobility restrictions during the COVID-19 pandemics.

2.1 Household Budget Survey

We derive data on the characteristics and expenditures of the Italian households from the Household Budget Survey (HBS) of the National Institute of Statistics (ISTAT). The HBS are the unique and official surveys on the expenditures of households living in Italy, providing also the reference for the official estimates of relative and absolute poverty in Italy, and the measure of inflation by household expenditure classes. The unit of analysis is the de facto household, and the data – collected annually – cover the entire Italian territory.¹ The HBS provides detailed monthly information on expenditures and household characteristics for approximately 20,000 households per year. For our analysis, we use monthly data from 2019 and 2020, thus spanning the period before and after the onset of the pandemic. Our sample includes 18718 households surveyed in 2019 and 25668 in 2020, which, once weighted using official sampling weights, represent the population of Italian households – 25994468 in 2019 and 26079006 in 2020. Table 1 presents the complete set of summary statistics for both households and household heads. On average, households in the sample consist of 2.3 members and reside in dwellings of about 97 square meters. With respect to the household heads, on average, they contribute for the great part of the family income (roughly 80%), and they are mostly males ($\approx 64\%$) and native ($\approx 90\%$) of 58 years age old, married ($\approx 50\%$) with a middle or high education ($\approx 65\%$), employed ($\approx 50\%$) mostly in manufacturing ($\approx 20\%$) or public, education and health service sectors ($\approx 18\%$).

Relevantly, to the best of our knowledge, we are the first to obtain access to highly detailed municipality-level data, which allow us to capture variation across house-

¹The statistical unit is the de facto household, defined as a group of cohabiting individuals connected by marriage, kinship, affinity, adoption, guardianship, or emotional ties, who share household expenses and/or income.

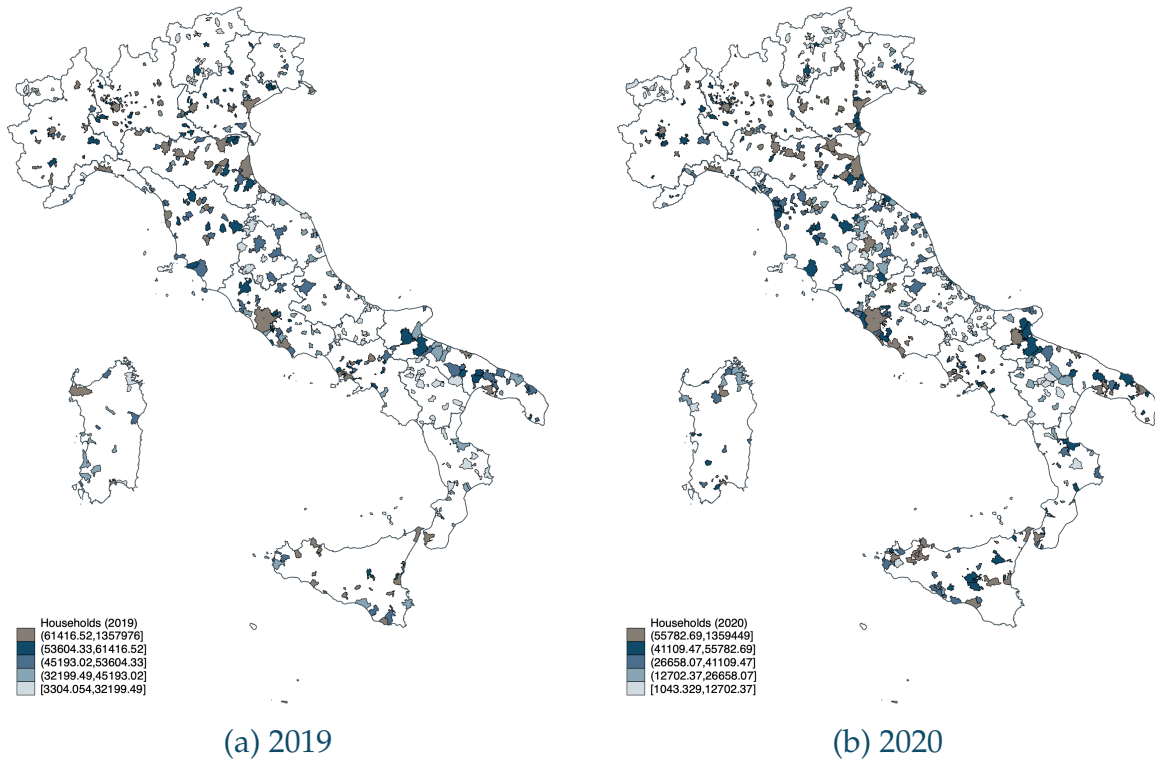
TABLE 1: SUMMARY CHARACTERISTICS OF HOUSEHOLDS AND HOUSEHOLDS' HEAD

	2019					2020				
	count	mean	sd	min	max	count	mean	sd	min	max
Households' characteristics										
Household income (decile)	18718	5.500	2.872	1	10	25668	5.500	2.872	1	10
Number components household	18718	2.306	1.252	1	9	25668	2.292	1.233	1	12.000
Households' head characteristics										
Male	18718	0.643	0.479	0	1	25668	0.632	0.482	0	1
Native	18718	0.898	0.302	0	1	25668	0.897	0.303	0	1
Age	18718	58.23	16.60	19	109	25668	58.35	16.60	18	104
Single	18718	0.206	0.404	0	1	25668	0.213	0.409	0	1
Married	18718	0.487	0.500	0	1	25668	0.484	0.500	0	1
Separated	18718	0.135	0.341	0	1	25668	0.136	0.342	0	1
Widow	18718	0.172	0.378	0	1	25668	0.167	0.373	0	1
No education - Elementary	18718	0.200	0.400	0	1	25668	0.188	0.391	0	1
Middle education	18718	0.311	0.463	0	1	25668	0.311	0.463	0	1
High education	18718	0.338	0.473	0	1	25668	0.346	0.476	0	1
Degree	18718	0.138	0.345	0	1	25668	0.141	0.348	0	1
Advanced education	18718	0.005	0.068	0	1	25668	0.015	0.121	0	1
Employed	18718	0.507	0.500	0	1	25668	0.493	0.500	0	1
Tenured (tempo indeterminato)	18718	0.342	0.475	0	1	25668	0.335	0.472	0	1
Agriculture and Extractive Industries	16994	0.074	0.262	0	1	23497	0.073	0.260	0	1
Manufacturing	16994	0.215	0.411	0	1	23497	0.196	0.397	0	1
Energy and Environmental Services (D-E)	16994	0.016	0.124	0	1	23497	0.018	0.131	0	1
Construction (F)	16994	0.077	0.266	0	1	23497	0.073	0.261	0	1
Trade and Repair Services (G)	16994	0.120	0.326	0	1	23497	0.126	0.332	0	1
Transport and Logistics (H)	16994	0.053	0.225	0	1	23497	0.052	0.222	0	1
Accommodation and Food Services (I)	16994	0.040	0.196	0	1	23497	0.045	0.208	0	1
Information and Communication Technologies (J)	16994	0.023	0.148	0	1	23497	0.025	0.155	0	1
Financial and Real Estate Activities (K-L)	16994	0.031	0.173	0	1	23497	0.036	0.185	0	1
Professional and Business Services (M-N)	16994	0.053	0.225	0	1	23497	0.058	0.233	0	1
Public, Education and Health Services (O-Q)	16994	0.179	0.384	0	1	23497	0.180	0.384	0	1
Cultural, Recreational, Household and Other Services (R-U)	16994	0.119	0.323	0	1	23497	0.120	0.325	0	1
Income (%)	18066	79.937	25.144	0	100	24770	79.778	25.164	0	100

holds within the smallest administrative units in Italy. In our data, there are 485 and 622 municipalities in the years 2019 and 2020, respectively, spanning the whole national territory (Figure 1).

Since 2014, the survey, based on the harmonized international classification of expenditure voices (Classification of Individual COnsumption by Purpose – Coicop), records all expenses incurred by resident households to purchase goods and services for final consumption, including also self-consumed goods, imputed rents, and gifts. Any expenditure not related to consumption such as the purchase of houses or land, payment of taxes, or professional expenses is excluded from the survey. In particular, the survey records monthly expenditures of the households at current Euros for roughly 488 items at 5-digits level, which are then aggregated into 12 chapters at 2-digits level: food and non-alcoholic beverages; alcoholic beverages and tobacco; clothing and footwear; housing, water, electricity, gas and other fuels (including renovation work); furniture, household goods and services; health; transport; information and communication; recreation, sport and culture; education; restaurants and accommodation; other goods and services, including insurance and financial services, personal

FIGURE 1: GEOGRAPHICAL DISTRIBUTION OF HBS HOUSEHOLDS



care, social protection services.² Figure 2 provides an illustrative example of the item-level disaggregation for a subset of the first chapter.

To test the hypotheses described above, we follow Charles et al. (2009), De Giorgi et al. (2020), and Heffetz (2011) who proxy for the degree of conspicuous consumption with the visibility of the goods and services consumed. Heffetz (2011) firstly noted that the higher the visibility of the items, the higher their conspicuousness. Likewise, Charles et al. (2009) and De Giorgi et al. (2020) develop classifications of “visible goods” to study the extent to which the consumption pattern of the households is shaped by status-seeking motives (as in conspicuous consumption models), defining “visible consumption” as the “households’ consumption of items that are readily observable in anonymous social interactions” (Charles et al., 2009). At this end, we begin by aggregating 429 out of the 488 5-digits expenditure items into the 16 categories illustrated in Table 2. In addition, we also use the total expenditures of each household to proxy for the permanent income as the HBS do not report household exact income but only its decile in the distribution.

While for most goods the definition of visibility is quite clear (i.e. jewelry versus

²The recorded amount refers to the total expenditure at the time of purchase, regardless of when the good or service is actually consumed or how it is paid for.

FIGURE 2: HBS EXPENDITURE AND ITEMS DISAGGREGATION

Ecoclip Chapter (2 digit)	Ecoclip Chapter Label	Ecoclip Group (3-digit)	Ecoclip Group Label	Ecoclip Class (4-digit)	Ecoclip Class Label	Ecoclip Subclass (5-digit)	Ecoclip Subclass Label	Expense Variables	Expense Variable Labels (ITA)
g_01	Food products and non-alcoholic beverages	c_01_1	Food Products	b_01_1_1	Bread and cereals	a_01_1_1_1	Rice	d004_mensile	'Riso'
						a_01_1_1_2	Flour and other cereals	d005_mensile	'Orzo, farro e altri cereali in chicchi'
						a_01_1_1_3	Bread	d010_mensile	'Farina di cereali'
						a_01_1_1_4	Other bakery and pastry products	d001_mensile	'Pane'
								d002_mensile	'Grisini e cracker'
								d007_mensile	'Biscotti, merendine e altra pasticceria'
								d011_mensile	'Altri stuzzichini salati'
						a_01_1_1_5	Pizza and quiches	d006_mensile	'Torte, pizze, dolci tipici'
						a_01_1_1_6	Dry pasta, fresh pasta and pasta preparations	d003_mensile	'Pasta'
						a_01_1_1_7	Breakfast cereals	d008_mensile	'Cereali da prima colazione'
						a_01_1_1_8	Other cereal-based products	d072_mensile	'Basi per pizze, focacce e torte'
				b_01_1_2	Meats	a_01_1_2_1	Beef	d022_mensile	'Carne bovina'
						a_01_1_2_2	Pork	d023_mensile	'Carne suina'
						a_01_1_2_3	Sheep and goat meat	d024_mensile	'Carne ovina'
						a_01_1_2_4	Poultry	d025_mensile	'Pollame'
						a_01_1_2_5	Other meats	d026_mensile	'Altre carni'
						a_01_1_2_6	Offal	d027_mensile	'Frittatine'
						a_01_1_2_7	Cured meats	d021_mensile	'Salumi'
						a_01_1_2_9	Other meat-based preparations	d028_mensile	'Altri misti di carne. Carne in scatola e altri preparati a base di carne'
				b_01_1_3	Fish and seafood products	a_01_1_3_1	Fresh or chilled fish	d029_mensile	'Pesce fresco'
						a_01_1_3_2	Frozen fish	d031_mensile	'Pesce surgelato'
						a_01_1_3_3	Fresh or chilled seafood	d030_mensile	'Frutti di mare, molluschi e crostacei freschi'
						a_01_1_3_4	Frozen seafood	d032_mensile	'Frutti di mare, molluschi e crostacei surgelati'
						a_01_1_3_5	Dried, smoked or salted fish and seafood	d033_mensile	'Prodotti essiccati, affumicati, sotto sale'
						a_01_1_3_6	Other preserved or processed fish and seafood	d034_mensile	'Prodotti in scatola e altri preparati di pesce, frutti di mare, molluschi, crostacei'

home insurance), we follow Charles et al. (2009) and De Giorgi et al. (2020) to classify all the expenditure items according to their potential nature of visible consumption. We build three versions of the “visible goods” measures as illustrated in Table 3. The first two measures are based on the classification supplied by Charles et al. (2009). We create two versions of this classification; the first one, “strict”, including only clothing & jewelry, personal care items, and new cars or motorcycles, and the other, “full”, that adds to the strict measure other items as reported also in Charles et al. (2009). For robustness, we also test with another classification employed by De Giorgi et al. (2020). In the main analysis of the paper, we use the full measure as derived following Charles et al. (2009).

Table 4 presents the summary statistics for the total expenditures as well as for the expenditures of each category and visible goods items.

In order to assuage concerns that our approach and definitions of visible consumption are actually able to match the consumption in luxury goods or, more broadly, conspicuous consumption, we begin by examining the share of the overall budget allocated to visible goods. Figure 3 plots the estimated nonlinear Engel curves for visible expenditures separately for the 2019 and 2020 years. To construct these curves, we regress log visible expenditures on log total expenditures, without or with city and month fixed effects (panel a and b, respectively). The figure indicates that visible goods behave as luxury items at every level of log total expenditures with an elasticity much greater than one and increasing along the income (i.e., expenditures) distribution as one would expect in a status-seeking model with conspicuous consumption (Heffetz, 2011).

Finally, Figures 4 and 5 display the monthly trends in (log) visible consumption

TABLE 2: EXPENDITURES AND GOODS: DEFINITION AND CLASSIFICATION

Category	Description
Total expenditure	Household total expenditures
Clothing & Jewelry	Clothing, jewelry, watches (incl. reparations)
Personal care	Hair salon / barber shop (men, women, children), beauty treatments, personal care and hygiene products, equipment (incl. reparation)
Cars/Moto (new)	Cars & motorcycles (new)
Vehicle (new)	Cars and motorcycles, campers, hang gliders, boats, canoes
Vehicle (extended)	Cars & motorcycles (new & used), tires, vehicle spare parts, accessories (satellite navigation, hi-fi systems, ...)
Housing	Real rental income, imputed rents, repairs and maintenance, utilities (energy, gas...)
Food	Food
Alcohol & Tobacco	Alcohol & tobacco
Mobile phones	Phones (incl. repair)
Transport (other)	Diesel, gasoline, other fuels, hourly parking, tolls, train, tram/subway, bus, taxi, ship/ferry, car/motorcycle insurance
Furnishing	Furniture and furnishings, carpets and other floor coverings, home textiles, household appliances and devices, glassware, tableware, and household utensils, home and garden tools and equipment, goods and services for routine household maintenance
Education	Expenditures for all levels
Health	All (incl. insurances)
Recreation outside (restaurants)	Hospitality and catering services (restaurants, hotels)
Entertainment private	Recreation, entertainment, and culture (excl. campers, hang gliders, boats, horses, canoes, sporting events, cultural and naturalistic visits, shows, day trips in Italy and abroad)
Entertainment public	Sporting events, cultural and naturalistic visits, shows, day trips in Italy and abroad

and in its share of total expenditures for 2019 and 2020. The figures indicate that, although the (log) level of visible consumption experienced only a modest decline in March and April 2020 following the outbreak of COVID-19, the evolution of its share in total spending shows a much sharper adjustment. In particular, the visible-goods share nearly halved in March 2020 relative to the same month in 2019, falling from about 10% to roughly 5% when the pandemic struck. Moreover, throughout 2020 the share of visible goods remained well below its 2019 counterpart, particularly during

TABLE 3: VISIBLE GOODS: DEFINITION AND CLASSIFICATION

Variable	Description
Visible goods (strict, Charles et al.)	Clothing & Jewelry, Personal care, Cars/Moto (new)
Visible goods (full, Charles et al.)	+ Cars & motorcycles (new & used), tires, vehicle spare parts, accessories (satellite navigation, hi-fi systems, ...)
Visible goods (De Giorgi et al.)	Clothing & Jewelry, Personal care, Vehicle (extended), Alcohol & tobacco, Entertainment public, Recreation outside

TABLE 4: SUMMARY STATISTICS: HOUSEHOLDS EXPENDITURES

	2019				2020			
	mean	sd	min	max	mean	sd	min	max
Total expenditure	2559.851	1619.897	110.000	34706.441	2328.232	1469.758	75.000	22375.699
Clothing & Jewelry	120.591	218.792	0.000	14632.550	91.772	187.730	0	4316.000
Personal care	74.889	103.455	0	1156.060	61.038	96.109	0	1402.530
Cars/motorcycles (new)	58.678	308.842	0	4166.670	50.436	296.948	0	7666.670
Vehicles (new)	60.035	314.400	0	4166.670	51.798	301.477	0	7666.670
Vehicles (new/used/parts/garage)	122.522	395.678	0	6250	101.913	357.579	0	7666.670
Housing (including utilities)	896.049	585.939	48.000	10216.580	893.214	602.963	42.500	13906.760
Food	464.272	328.256	0	3980.830	467.557	324.117	0	4719.490
Alcohol and tobacco	46.096	76.636	0	1171.570	42.542	75.100	0	1085.000
Mobile phones	5.343	14.904	0	250	5.049	15.041	0	258.330
Other transport	192.837	180.020	0	2311.660	148.175	146.110	0	1759.600
Household furnishing	109.969	287.269	0	26798.301	103.661	280.865	0	11118.940
Education	15.826	68.863	0	2916.670	13.633	63.190	0	2916.670
Health	121.387	204.516	0	5047.960	111.059	193.999	0	4780.410
Entertainment outside (restaurants)	129.978	212.766	0	3520.320	79.407	164.211	0	2527.250
Entertainment private	114.977	203.248	0	4550.670	88.728	163.384	0	3436.500
Entertainment public	9.131	42.681	0	2597.490	2.727	31.927	0	2211.850
Visible goods								
Visible goods (strict)	254.158	429.560	0	15139.609	203.246	390.149	0	7887.900
Visible goods (full)	318.002	508.740	0	15139.609	254.723	450.730	0	7887.900
Visible goods (Pistaferri)	503.207	633.509	0	16122.060	379.398	543.188	0	8753.480

the months subject to the strictest pandemic-related restrictions – namely, from March to July and again from October through December.

FIGURE 3: ENGEL CURVES

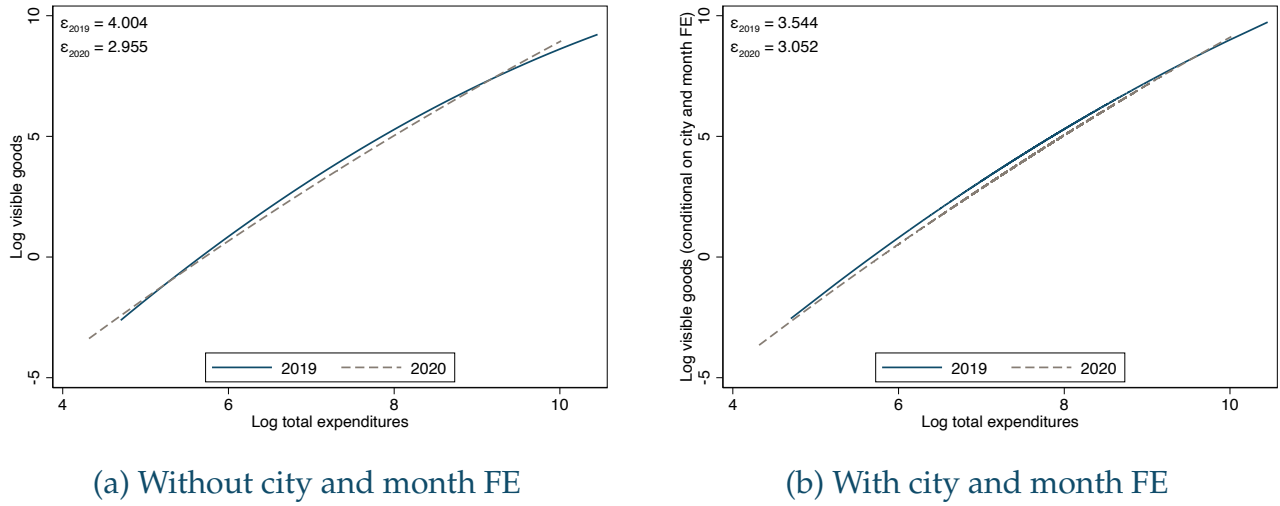


FIGURE 4: (LOG OF) VISIBLE CONSUMPTION, 2019–2020

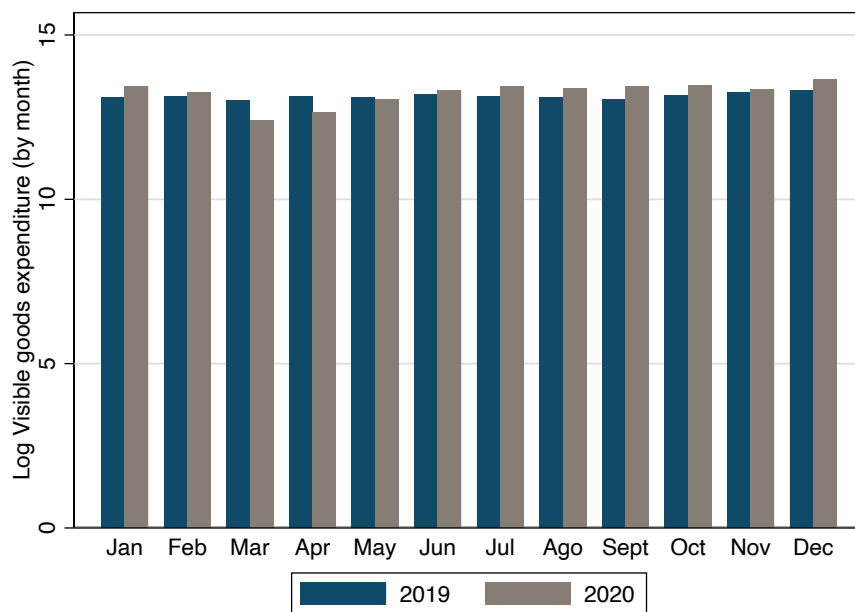


FIGURE 5: SHARE OF VISIBLE CONSUMPTION, 2019–2020

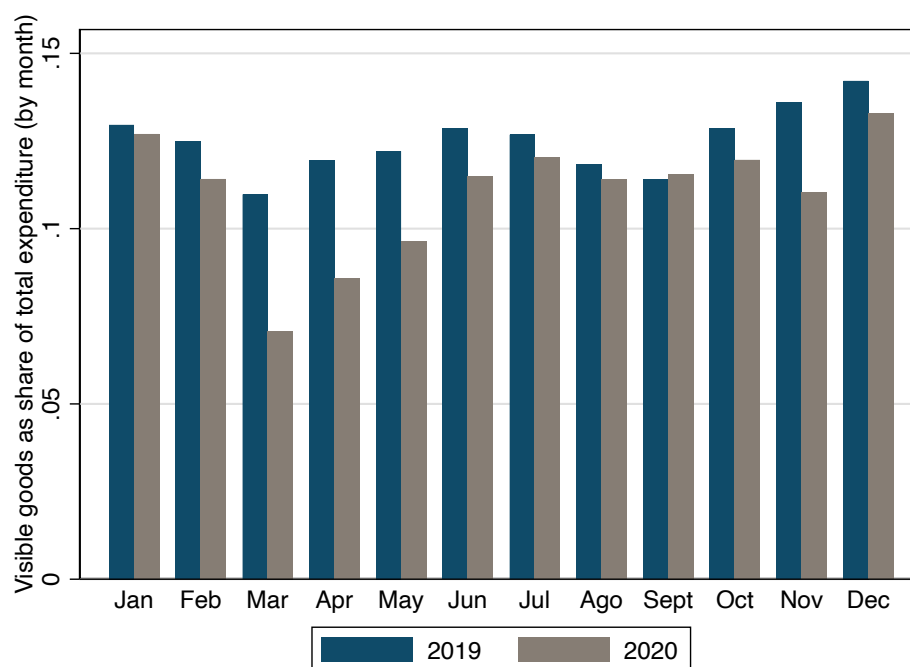
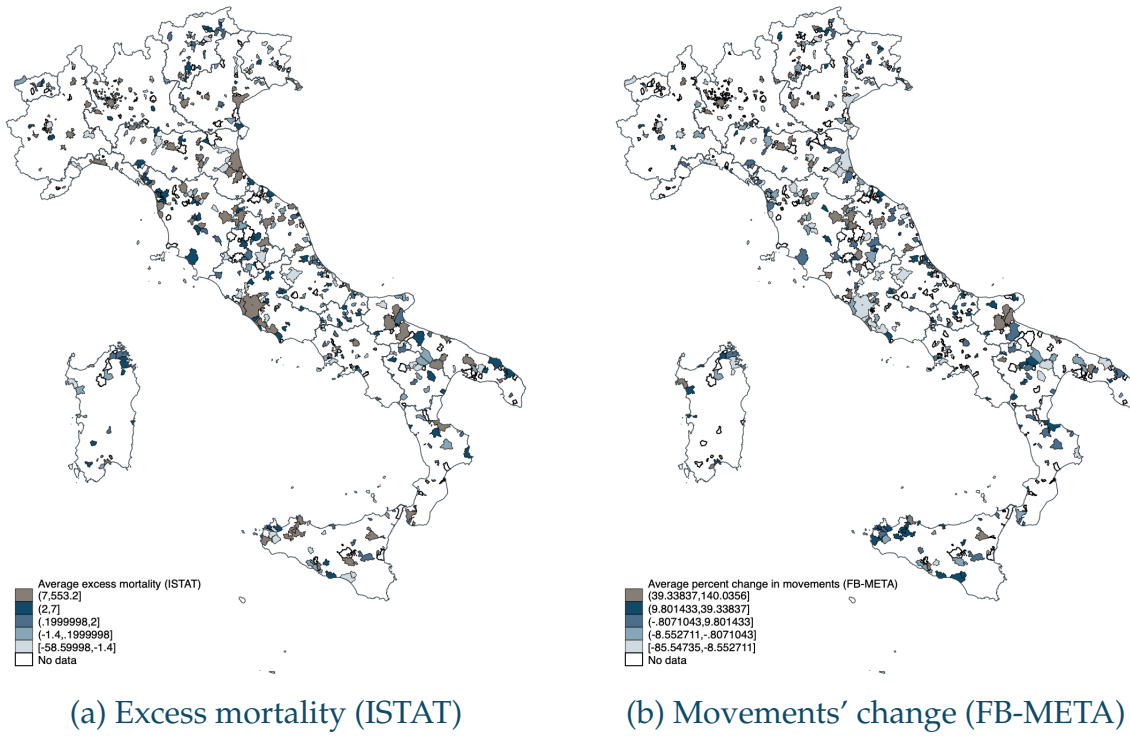


FIGURE 6: COVID-19 INTENSITY AND MOVEMENTS CHANGE

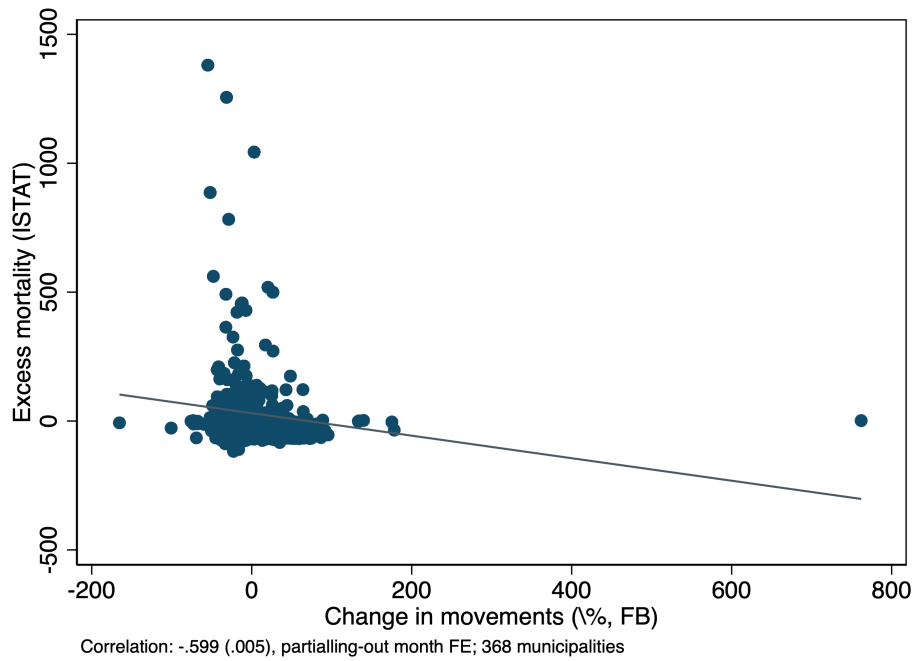


2.2 COVID-19 and mobility data

The second source of data we employ is meant to measure the reduction in social interactions due to the restrictions generated by the onset of COVID-19 pandemics.

In response to the COVID-19 outbreak, the Italian government introduced a first nationwide lock-down on March 11, 2020, and strengthened these measures on March 22 with a second, more stringent economic shutdown. During this initial phase, many public-facing activities – such as restaurants, gyms, and a wide range of retail services – were suspended. The restrictions tightened further with the closure of all non-essential economic sectors and severe limitations on individual mobility, except for certified work or health-related reasons. After this early period, the intensity of the restrictions increasingly depended on the local evolution of the pandemic. Regional authorities were granted discretion to determine the level of closures and, consequently, the degree of permitted social interaction at the municipal level. Because highly granular data capturing local variation in the reduction of social interactions are not available, we rely instead on municipal monthly COVID-19 mortality as a proxy for the local severity of the pandemic and hence the reduction of social interactions. Following the approach adopted in the existing literature and suggested by public health authorities (Borri et al., 2021), we measure COVID-19 mortality using excess deaths,

FIGURE 7: VALIDATION



defined as the difference between the number of monthly deaths in 2020 and the average number of deaths occurring on the same month in the same municipality over the years 2015-2019. Using official ISTAT municipal-level mortality records for 2015-2020, we then construct a monthly measure of excess mortality for each municipality as the difference between total deaths in each month of 2020 and the corresponding 2015-2019 average. Panel (a) in Figure 6 shows the variability in the 2020 average excess mortality across all the municipalities in the HBS sample.

To validate this approach and interpretation, we also employ a more direct measure that, however, is available only for a subset of municipalities of the HBS. In particular, we obtained data on the mobility of individuals from the Facebook Movement Range dataset, produced within its META's Data For Good program and Disease Prevention Maps, covering movements in and out a sample of municipalities during the months between March and December 2020.³ This dataset is constructed from aggregated and anonymized location information of Facebook users who opt into sharing their location history. Although the sample is not fully representative of the entire population, prior research has shown that these data provide a timely and spatially detailed proxy for local mobility dynamics. Facebook aggregates users' movements across small geographic tiles and maps them onto administrative boundaries, thereby

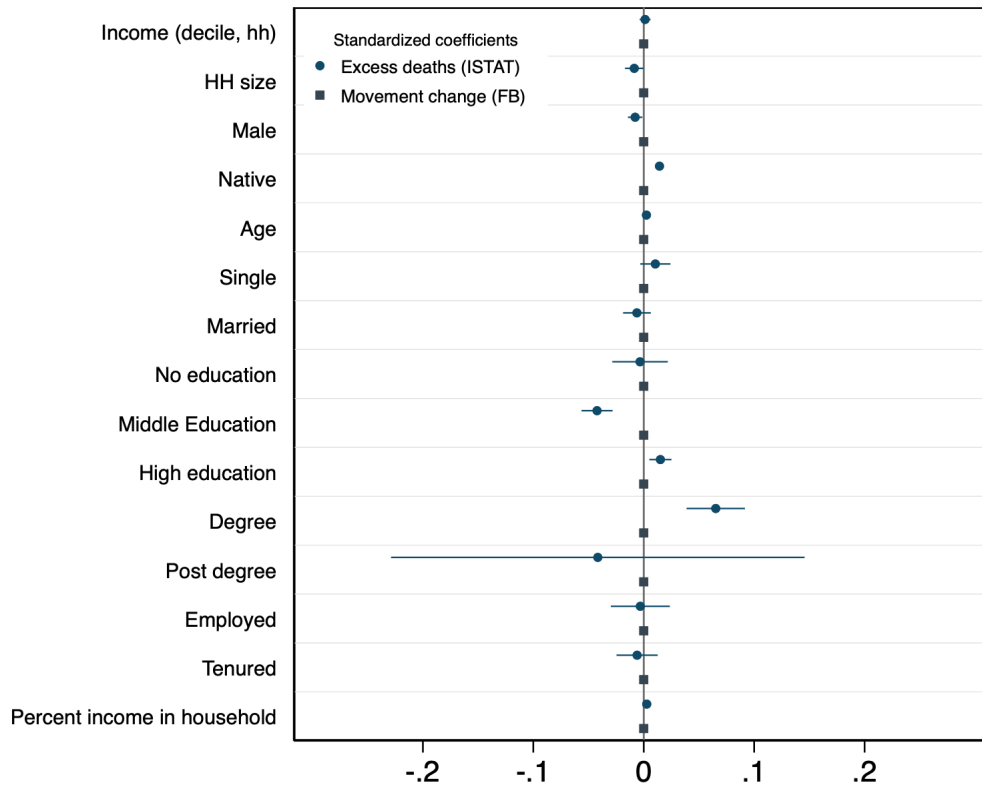
³We greatly thank Angelo Facchini for giving us access and help for the management of these data.

offering consistent measures of mobility across all Italian municipalities. A key feature of the Movement Range data is that mobility is expressed relative to a pre-pandemic baseline. For each day, Facebook computes how much users move compared with their typical behavior before COVID-19, using as a reference period the corresponding days of the previous year. The metric therefore captures the percentage change in movement with respect to normal, pre-COVID patterns, allowing us to interpret mobility levels as increases or reductions compared to usual conditions rather than as absolute flows. This baseline adjustment is crucial during the pandemic, as it isolates the extent to which mobility declined in response to restrictions and behavioral adjustments. We used the movements in and out of each municipality, aggregated to the monthly level. This measure summarizes all within and inter-municipal movements observed in the dataset and reflects the intensity of mobility for each locality. By relying on a baseline-adjusted metric, it offers a coherent indicator of how strongly individuals reduced or modified their mobility during different phases of the pandemic and provides a meaningful comparison of mobility patterns across municipalities and over time. Panel (b) in Figure 6 shows the variation in the 2020 average percent change in movements across a subsample of municipalities.

Unfortunately, the Facebook data has a smaller coverage in terms of municipalities with respect to the municipalities in the HBS sample. In fact, because of privacy concerns, Facebook does not report data when the number of individual users is smaller than 10. In particular, using these data we would lose 205 out of 622 municipalities corresponding to roughly 40% of the observations. For this reason, and what said above about the correlation between mortality intensity and interactions, we choose to use the monthly excess mortality as a proxy of the reduction in social interactions. Figure 7 presents the scatterplot of the two measures showing, indeed, the strong and negative correlation between excess mortality and the percent movement change across municipalities.

Finally, Figure 8 presents a balance check of the two measures against the individual characteristics of the households' heads. We regress each individual characteristics on the y-axis against each measure separately, conditional on city and month fixed effects. Due to the unexpected nature of the COVID-19 pandemics, it is reassuring that, overall, there are no particular individuals' characteristics correlated with the intensity of the pandemics and the associated reduction in movements that can also be correlated with the consumption patterns.

FIGURE 8: BALANCE CHECK



3 Empirical Analysis

We study the effect of the COVID-19 pandemics and the associated reduction in social interactions on the pattern of consumption and visible expenditures in three stages using fixed-effects and event-study empirical specifications. We start by looking at the empirical relation between the intensity of the COVID-19 and the pattern of consumption in visible goods across municipalities. Then, we proceed by empirically showing that the expenditures in visible goods well approximate the behavior of the households towards positional consumption and conspicuous consumption. Finally, we put together the two pieces of evidence and study the relation between expenditures in visible goods, COVID-19 intensity and status-seeking behavior.

3.1 Visible expenditures and COVID-19

We begin by examining how the local severity of the COVID-19 shock shaped, on average, household consumption patterns by estimating the specification in Equation (1) that relates monthly visible expenditures to excess mortality at the municipality

level.

$$\log(\text{Visible})_{icm} = \beta \text{ExcessMortality}_{cm} + \gamma \log(\text{Total Expenditures})_{icm} + \mu_c + \delta_m + \mathbf{X}'_{icm}\theta + \varepsilon_{icm} \quad (1)$$

The Equation (1) exploits within-municipality, over-time variation by including both municipality (μ_c) and month (δ_m) fixed effects, thereby purging the estimates of long-run differences in local economic conditions as well as nationwide contemporaneous shocks. The main treatment in this first part of the analysis is the variable *Excess mortality*, defined in the above section as the deviation of deaths in each month of 2020 from the average recorded in the same month over the years 2015–2019, re-scaled $\times 1000$ households in the specifications. Hence, this first analysis exploits only variation over the year 2020. Because the data are repeated cross-sections and household composition changes across months, the preferred specification additionally includes a set of individual controls that may affect the consumption of visible goods. In particular, we control for the full set of variables presented in Table 1 capturing the characteristics of the households and the households' heads. Specifically, as we do not have the whole income of the households, we control for the household income decile in addition to the number of components of each household. We also for the individual characteristics of the household heads to control for omitted time-varying factor that may affect the household consumption patterns. In particular, we add gender (dummy *Male*), immigrant status (dummy *Native*), *Age*, marriage status with the dummies *Single*, *Married*, *Separated* and *Widowed*, the education level with the dummies *No*, *Middle*, *High*, *BA*, and post-BA, employment status (dummy *Employed*), sector of employment as reported in descriptive statistics, and the percent of household income apportioned by the households head. Finally, we control for the (log) total expenditures. Indeed, we have only poor information on the income of the households, their decile, and of the household heads, their percent contribution in the overall household income. Hence, to appropriately account for a permanent income factor that may affect the inter-temporal consumption choices of the households, we follow Charles et al. (2009) and De Giorgi et al. (2020) in using the total expenditures as a proxy for the total – permanent – income of the households.

The results, presented in Table 5, show a clear and robust negative relationship between excess mortality and spending on visible goods. Across specifications, the estimated coefficient is sizeable and statistically significant, with magnitudes around

TABLE 5: VISIBLE EXPENDITURES AND COVID-19: AVERAGE EFFECTS

	Log Visible goods (full)		
	(1)	(2)	(3)
Excess mortality (x1000)	-0.138** (0.066)	-0.169*** (0.055)	-0.181*** (0.060)
Log total expenditures		2.188*** (0.030)	2.204*** (0.038)
Municipality FE	yes	yes	yes
Month FE	yes	yes	yes
Individual controls	no	no	yes
R^2	0.110	0.422	0.442
Observations	25668	25668	22901

Notes. Robust standard errors clustered at municipality level in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

−0.18 in the preferred model implying . This indicates that, holding total expenditures constant, higher local mortality is associated with a marked contraction in visible consumption. The inclusion of log total expenditures, whose coefficient is strongly positive and above two, highlights the discretionary nature of visible goods, yet the mortality coefficient remains large, suggesting a mechanism that operates beyond income or budget effects. A plausible interpretation is that the social and psychological environment in high-mortality periods reduces the value of conspicuous consumption, either through diminished opportunities for social interaction or through heightened uncertainty that induces a reallocation toward less discretionary categories. Taken together, the evidence shows that mortality shocks during the pandemic reshaped not only the level but also the composition of household spending, consistent with theories of visible consumption and behavioral responses to risk.

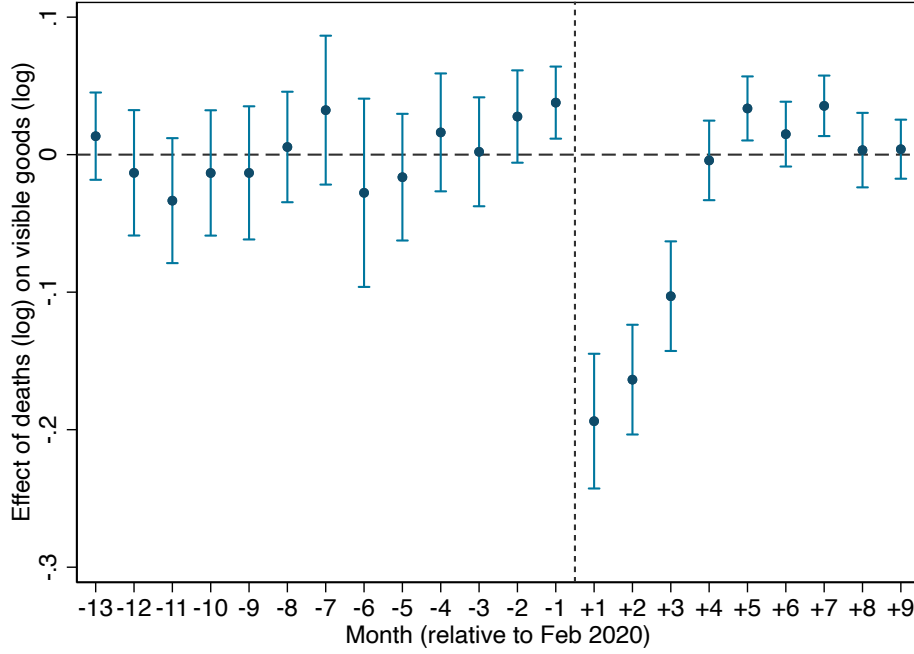
To complement the average effects estimated above and to capture the dynamic response of visible consumption to pandemic severity, we estimate an event-study specification that interacts the monthly number of deaths in each municipality with a series of month dummies centered on February 2020. Formally, we estimate Equation (2) regressing the logarithm of visible expenditures on the interaction between the (log) number of deaths in each municipality and month between January 2019 and December 2020 ($\log(Deaths_{cm})$), and the indicators $D_{m=k}$ for each month $m = k$ before and after the onset of the pandemic, excluding February 2020 as the reference month. This strategy relies on the same within-municipality variation as the baseline model, but by

allowing the coefficient on deaths to vary flexibly over time, it enables us to trace the evolution of the mortality-consumption relationship both in the pre-pandemic period and throughout the subsequent phases of the health crisis. Municipality fixed effects control for time-invariant local characteristics, while contemporaneous total household expenditure is included to net out changes in the household budget and to isolate shifts in the allocation of spending toward visible goods. The identifying assumption is the standard one for event-study designs: in the absence of the pandemic shock, the association between local mortality and visible consumption would not have displayed differential trends across months, implying that pre-2020 coefficients should remain close to zero.

$$\begin{aligned} \log(Visible)_{icm} = & \sum_{k \neq 0} \beta_k [D_{m=k} \times \log(Deaths_{cm})] + \\ & + \gamma \log(Total\ Expenditure)_{icm} + \mu_c + \varepsilon_{icm} \end{aligned} \quad (2)$$

The estimated dynamic coefficients, plotted in Figure 9, strongly support this interpretation. Prior to February 2020, the estimated effects are small, statistically indistinguishable from zero, and display no discernible pattern, suggesting that mortality fluctuations in normal times bear no systematic relationship with visible consumption and validating the parallel-trends condition. Starting in March 2020, however, the coefficients drop sharply, indicating that increases in local deaths are associated with a sizeable reduction in spending on visible goods. The trough of this decline corresponds to the peak of the first pandemic wave, where the estimated elasticity becomes markedly negative. In the subsequent months, the coefficients gradually recover toward zero, though they remain weaker than in the pre-pandemic period, pointing to a partial but incomplete normalization of consumer behavior. This dynamic profile is consistent with a mechanism in which visible consumption contracts most strongly during periods of acute health risk and then progressively rebounds as mortality pressure subsides and social interactions resume. Taken together, the event-study evidence reinforces the findings from the baseline specification: mortality shocks exert a distinct and time-varying influence on the composition of household spending, with visible goods being particularly sensitive to heightened uncertainty and reduced opportunities for social display.

FIGURE 9: EVENT-STUDY COEFFICIENTS: VISIBLE CONSUMPTION AND COVID-19



3.2 Visible expenditures as positional consumption

In this section, we explore whether the consumption of visible goods responds not only to individual budget constraints and local mortality shocks but also to relative consumption considerations. At this end, we estimate the specification in Eq. (3) that incorporates a measure of reference-group spending. The motivation is rooted in theories of status-seeking and positional consumption, which predict that demand for conspicuous goods is shaped by the consumption norms of one's local peer group. Formally, in Eq. (3) we regress the log household-level visible expenditures on the logarithm of average consumption in the municipality-month, allowing this reference measure to operate both directly and interactively with excess mortality. Following the relevant literature (Charles et al., 2009; De Giorgi et al., 2020), we construct the reference-group consumption in two alternative ways: first, as the logarithm of average total expenditures in the city-month, and second as the logarithm of average visible expenditures in the same city-month. These two constructions allow us to distinguish whether visible spending responds to general consumption norms or specifically to local patterns of conspicuous consumption. The model further includes municipality and month fixed effects, as well as household controls and total expenditures where indicated, ensuring that the estimates reflect within-city variation across time after netting out local structural differences and aggregate monthly shocks.

$$\begin{aligned}
\log(\text{Visible})_{icm} = & \alpha \log(\text{Reference consumption})_{cm} + \beta \text{Excess Mortality}_{cm} + \\
& + \delta [\log(\text{Reference consumption})_{cm} \times \text{Excess Mortality}_{cm}] + \quad (3) \\
& + \mathbf{X}'_{icm} \theta + \mu_c + \lambda_m + \varepsilon_{icm}
\end{aligned}$$

The results reported in Table 6 reveal a strong and consistent role for reference-group consumption. Across specifications, the coefficient on the reference measure is positive and significant, indicating that visible expenditures rise when average consumption in the municipality increases. The effect is substantially larger when the reference measure is based on visible spending itself, suggesting that households respond more intensely to peer-group norms in conspicuous categories than to general consumption patterns. This finding aligns closely with the idea that visible goods serve as positional signals whose value depends on the prevailing standards of the surrounding social environment. Turning to the mortality dimension, excess mortality retains a negative and statistically significant association with visible consumption in the baseline interaction-free specification, consistent with earlier results. More importantly, the interaction term between excess mortality and reference consumption is negative and significant when the reference group is defined in terms of visible expenditures. This pattern indicates that positional pressures weaken in high-mortality contexts: when social interactions contract and the opportunities for signalling diminish, the sensitivity of visible consumption to local spending norms declines. When the reference group is based on total consumption, the interaction coefficient is negative but imprecisely estimated, suggesting that the erosion of positional behaviour during the pandemic is particularly salient for conspicuous categories. Finally, the coefficient on household total expenditures remains large and stable across specifications, confirming the highly elastic nature of visible spending. Overall, these results show that visible consumption is shaped jointly by status-seeking forces and mortality-induced behavioural adjustments, and that the salience of positional motives declines precisely when the pandemic most severely restricts opportunities for social display.

TABLE 6: VISIBLE EXPENDITURES AS POSITIONAL CONSUMPTION

	Dependent: Log Visible goods					
	Reference group (City-month log average consumption)					
	Total consumption			Visible consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Log average consumption	0.125*** (0.048)	0.129* (0.071)	0.136* (0.072)	0.451*** (0.015)	0.460*** (0.022)	0.463*** (0.022)
Excess mortality (x1000))		-0.185*** (0.059)	3.688 (3.489)		-0.204*** (0.044)	1.254*** (0.389)
Log average consumption $\times$ Excess mortality (x1000)			-0.489 (0.434)			-0.277*** (0.071)
Log Expenditure	2.167*** (0.030)	2.190*** (0.039)	2.190*** (0.039)	2.067*** (0.027)	2.099*** (0.037)	2.100*** (0.036)
Municipality FE	yes	yes	yes	yes	yes	yes
Month FE	yes	yes	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes	yes	yes
R^2	0.454	0.442	0.442	0.469	0.455	0.456
Observations	39452	22901	22901	39417	22866	22866

4 Conclusions

We provide evidence that visibility and social comparison drive a sizable share of household consumption. The COVID-19 shock, by removing opportunities for social display, led to a marked reduction in visible spending and attenuated peer effects. These findings suggest that conspicuous consumption is not a fixed trait but depends critically on the social environment. Future work will assess whether these behavioral adjustments persisted after restrictions eased.

References

- Abel, A. B. (1990). Asset prices under habit formation and catching up with the joneses. *American Economic Review Papers & Proceedings*, 80(2):38–42.
- Bachmann, R. et al. (2024). Consumption and saving during pandemics. *Journal of Monetary Economics*.
- Baker, S. et al. (2020). Income, liquidity, and consumption responses to covid-19. *Review of Financial Studies*, 33(11):5502–5551.
- Barro, R. J., Ursúa, J. F., and Weng, J. (2020). The coronavirus and the great influenza pandemic: Lessons from the “spanish flu” for the coronavirus’s potential effects on mortality and economic activity. NBER working paper 26866, National Bureau of Economic Research.

- Borri, N., Drago, F., Santantonio, C., and Sobbrío, F. (2021). The "Great Lockdown": Inactive workers and mortality by COVID-19. *Health economics*, 30(10):2367–2382.
- Charles, K., Hurst, E., and Roussanov, N. (2009). Conspicuous consumption and race. *Quarterly Journal of Economics*, 124(2):425–467.
- Christelis, D., Georgarakos, D., and Jappelli, T. (2020). The covid-19 crisis and consumption: Survey evidence from six eu countries. *ECB Working Paper Series*.
- De Giorgi, G., Frederiksen, A., and Pistaferri, L. (2020). Consumption networks. *Review of Economic Studies*, 87:130–163.
- Duesenberry, J. (1949). *Income, Saving, and the Theory of Consumer Behavior*. Harvard University Press.
- Dupor, B. and Liu, W.-F. (2003). Jealousy and equilibrium overconsumption. *American Economic Review*, 93(1):423–428.
- Frank, R. (1985). The demand for unobservable and other nonpositional goods. *American Economic Review*, 75:101–116.
- Frank, R. H. (1999). *Luxury fever: Why money fails to satisfy in an era of excess*. Free Press, New York.
- Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics*, 93(4):1101–1117.
- Hirsch, F. (1976). *Social limits to growth*. Harvard, Cambridge: Massachusetts.
- Hopkins, E. and Kornienko, T. (2004). Running to keep in the same place: Consumer choice as a game of status. *American Economic Review*, 94(4):1085–1107.
- Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics*, 120(3):963–1002.
- Rayo, L. and Becker, G. S. (2007). Evolutionary efficiency and happiness. *Journal of Political Economy*, 115(2):302–337.
- Robson, A. J. and Samuelson, L. (2011). The Evolutionary Foundations of Preferences. In Benhabib, J., Bisin, A., and Jackson, M., editors, *Handbook of Social Economics*, volume 1, chapter 7, pages 221–310. North-Holland, New York and Oxford.
- Veblen, T. (1899). *The Theory of the Leisure Class*. Macmillan.