Essays on Firm Behavior and Productivity

Andrea Petrella

TESI DOCTORAL UPF / ANY 2017

DIRECTOR DE LA TESI
Vasco M. Carvalho, University of Cambridge

Universitat Pompeu Fabra
Barcelona

Acknowledgments

My first thanks goes to Vasco Carvalho for having been an excellent professor, and a frank and insightful advisor. His guidance has been invaluable in shaping the way I think at economic questions. He taught me to be curious, to question my results and to look at problems from multiple angles. Not least, he had patience with me (and he needed a lot).

I would like to thank my friend and coauthor Andrea Linarello for having walked with me during all these years. His example taught me the importance of method and dedication. His company made method and dedication look fun. I have no difficulty in admitting that without his encouragement and help I would have hardly finished this thesis.

I owe a lot to my other colleagues and coauthors Antonio Accetturo, Andrea Lamorgese and Enrico Sette for sharing with me their knowledge and their time. Working with them has been instructive and stimulating, and I will always regard it as one of the luckiest circumstances in these years at the Bank of Italy.

Thanks to Marta, who gave me the motivation I needed to take the last steps. She lent me her determination and instilled me tranquility. Having her by my side is a daily demonstration that I can look with confidence to the future.

Thanks to Mapis and Ponti, for having been my home and my family in Barcelona. Our little routines and habits at Nou de Sant Francesc will always find a place in my most cherished memories. Thanks also to Ambro, Giulia and Lauretta, for enriching each day I spent in Barcelona.

A special thanks goes to Marta Araque, who helped me find a way through deadlines and administrative issues.

Thanks to Tesei, Commi, Aniolcito, Felipe, Paula, Italo, Christian, Bruno, Jagdish and all my fellow PhD students at UPF. Thanks to Lucia, Francesca, Sauro, Betta, Emanuele, Monica and all the colleagues at the Bank of Italy who shared with me their comments and suggestions.

Most importantly, I thank my parents, Valeria and Domenico, who are the cause and the aim of all of this.

Finally, thanks to all the *compadres*. You made my life better. Hopefully also a better person.

Abstract

This thesis investigates different aspects of firm behavior. In the first chapter I study if the quality of civil justice affects the firms' participation to Global Value Chains. I find that firms subject to less efficient courts are less likely to supply customized intermediate inputs to foreign firms. In the second chapter I analyze the impact of credit supply shocks on aggregate productivity. I find that a credit restriction depresses firm-level productivity growth. At the same time, aggregate productivity is sustained by the reallocation of resources towards more efficient firms. In the third chapter I study the determinants of the productivity advantage of firms located in urban areas. Results show that most of the urban productivity premium is explained by the sorting of more efficient firms to cities. The rest is explained by positive agglomeration externalities specific to each city.

Resum

Aquesta tesi investiga diferents aspectes del comportament d'empresa. En el primer capítol, estudio si la qualitat de la justícia civil afecta la participació de les empreses a les cadenes de valor global. Trobo que empreses subjectes a tribunals menys eficients tenen una menor probabilitat de subministrar productes intermediaris personalitzats a empreses estrangeres. En el segon capítol analitzo l'impacte dels xocs a l'oferta de crèdit en la productivitat agregada. Trobo que una restricció creditícia redueix el creixement de la productivitat a nivell d'empresa. Al mateix temps, la productivitat agregada se sosté per la reassignació de recursos cap a empreses més eficients. En el tercer capítol em centro en els determinants de l'avantatge de la productivitat de les empreses ubicades en zones urbanes. Els resultats mostren que la major part de l'avantatge de la productivitat urbana s'explica per la localització d'empreses més eficients a les ciutats. La resta s'explica per externalitats d'aglomeració positives específiques de cada ciutat.

Foreword

My doctoral thesis is a collection of three self-contained essays that study firm behavior and productivity in relation to local economic conditions and institutions.

In the first chapter, co-authored with Antonio Accetturo and Andrea Linarello, we study the relationship between the quality of civil justice and the firms' participation to Global Value Chains. Poor contract enforcement may raise the inherent riskiness of a contract, thus reducing the probability of firms to establish commercial relationships. We show that firms located in courts with longer trial lengths are less likely to supply customized intermediate inputs to foreign firms. Our empirical analyses exploit variations in the efficiency across Italian tribunals, and use a new dataset that records whether firms supply customized inputs abroad. Consistently with the theory, we find that the negative effect of judicial inefficiencies on the firms' participation to GVCs is stronger in industries producing goods requiring more relationship-specific investments. To confirm that the results are not influenced by omitted variables at the court level, we use a spatial regression discontinuity design that compares the probability of supplying customized inputs abroad for firms located on different sides of a court border, and are therefore characterized by different trial lengths.

The second chapter, co-authored with Andrea Linarello and Enrico Sette, focuses on the effect of credit supply shocks on aggregate productivity growth in the manufacturing sector. Using a dataset that covers the universe of Italian manufacturing firms, we distinguish between different components of aggregate productivity growth: within-firm productivity growth, the reallocation of labor across incumbent firms, and the extensive margins (entry and exit). Our findings show that a credit restriction on one hand depresses productivity growth at the firm level; on the other hand, it fosters the reallocation of employment shares from less to more productive firms. The more selective environment induced by a credit restriction also affects the extensive margins: it increases the contribution of exit, since more and relatively less productive firms leave the market, and further lowers the one of entry, since firms enter with a lower productivity when the credit availability is limited. We provide an aggregate quantification of the effects of the 2008–14 credit shrinkage, showing that the positive contribution of reallocation and exit has more than compensated the negative effect that credit crunch has had on within-firm productivity and entry.

In the third chapter, co-authored with Andrea Lamorgese, we study the determinants of the productivity advantage displayed by firms located in urban areas. Using detailed data on the universe of Italian firms, we are able

to disentangle the role played by two components: the sorting of intrinsically more productive firms to urban areas, and the extent of agglomeration economies specific to each city. While both channels significantly contribute to the urban productivity premium, we find that the former is more sizable than the latter. In order to shed some light on the mechanisms through which firms accrue a productivity advantage in a urban environment, we focus on firms relocating across cities. We set up a synthetic control exercise that builds appropriate counterfactuals to evaluate the productivity gains arising from relocation. The results show that firms reap the benefits of agglomeration in two ways: on one hand, the simple fact of being located in a city entails a static productivity premium; on the other hand, productivity gains accumulate more rapidly in urban areas, hinting at potentially faster learning processes for firms located in these environments.

Contents

List of Figures

List of Tables

Chapter 1

Legal Enforcement and Global Value Chains: Micro Evidence from Italian Manufacturing Firms

Joint with Antonio Accetturo and Andrea Linarello, Bank of Italy.

1.1 Introduction

Over the past twenty-five years, the ICT revolution, the steady lowering of trade barriers and transport costs, and the access to global markets by several low-wage countries have led to a remarkable structural change in the global economy (Antràs, 2015). The outcome is a new international division of labor, in which the production of final products is fragmented in Global Value Chains (GVC henceforth). The policy relevance for this structural change cannot be understated. UNCTAD (2013) shows a strong correlation between GDP growth and participation to GVCs, especially for developing countries; GVCs, in particular, are seen as an opportunity for small firms to access global trade, considering the presence of relevant fixed costs associated with the search of new buyers in large and unknown markets.

GVC are characterized by contractual relationships between intermediate producers and assemblers for the delivery of customized inputs that often require relationship-specific investments. When a country's ability to enforce contracts is weak, however, hold-up problems may arise. This, in turn, leads to suboptimal levels of investments and aggregate surplus (Antràs, 2003;

Antràs and Helpman, 2004), it affects the pattern of comparative advantage (Nunn, 2007; Levchenko, 2007), and it depresses international trade (Anderson and Marcouiller, 2002; Berkowitz et al., 2006). Despite in recent years the availability of micro data have substantially contributed to the improvement of our understanding of international trade, there is little or no empirical evidence on the relationship between the quality of institutions and participation to GVC using firm-level data. The lack of evidence is mostly driven by the fact that production fragmentation is difficult to measure at firm level; this is a major weakness of many empirical works, given the fact that most of the cross-country heterogeneity in the performance on international markets generally depend on firm-level characteristics.

In this paper we provide direct evidence that weak contract enforcement influences firm participation to GVC. We first present a theoretical model in which we show that the ability of a firm to supply an intermediate good to an international buyer crucially depends on the quality of domestic judiciary institutions; we also show that the negative effects of a bad contractual environment are amplified when the intermediate good needs complex contractual arrangements between the buyer and the seller. In the empirical part we focus on the case of Italy, which provides a good empirical setting. We build on a unique dataset from the 2011 Italian Census of Industry and Services, where we are able to identify manufacturing firms that supply customized inputs (within its group or to other companies) domestically or abroad. The Italian law codifies a specific contract type for the supply of customized inputs ("*contratto di subfornitura*, L. 192/1998"), in which the buyer provides design and production criteria to the supplier that performs the physical transformation activities. This contract is widely used in the Italian context (Lazerson, 1999). Because firms directly report whether they supply customized inputs under this particular contract arrangement, we are able to directly measure firms participation in GVC, improving over many existing measures of offshoring.

We exploit the variation in contract enforcement across courts within Italy and a specific feature of the Italian legal system to estimate the impact on firms' participation to GVCs. In case of litigation, the Italian law automatically determines the court in charge that corresponds to the one where the majority of the assets of the Italian firm are located. That court is also responsible for the direct enforcement of any decision for international disputes, made by either a foreign court or an international arbitration. In other words, there is no room for imposing a *choice-of-law* clause on the final enforcer. In 2011, Italy had 165 tribunal jurisdiction areas that display large differences in contract enforcement. According to the World Bank's *Doing business* figures, based on direct interviews with legal professionals, law enforcement in Italy is on average quite poor: Italy ranks 155th out of 185 countries. However, Italy also features a substantial variation in the quality of law enforcement between courts: as shown in Figure 1.1, in Bari it might take more than twice as long as in Turin to have a contract enforced. This dispersion in court efficiency is also confirmed by the data used in this paper, based on caseflow data provided by the Ministry of Justice: the difference in the average duration of a trial between the best- and the worst-performing tribunal is equal to 4.5 years.

Regression results show a strong negative correlation between judicial trial length in civil disputes and the firms' supply of customized inputs abroad. Firms located in courts with weaker institutions are less likely to participate in GVCs. The estimates suggest that a one-year increase in trial length is associated with a 1.9 percentage point reduction of this probability (about 1/10 of the unconditional probability, equal to 17.8%). This relationship should be stronger in those industries that produce goods that are characterized by relationship-specific investments. For each narrowly defined industry we measure relation specificity as the share of products that are not sold on organized markets according to the Rauch (1999) classification (following Nunn (2007)). We interact trial length with our measure of contract intensity in a Rajan-Zingales specification that includes court fixed effects. We show that the average effect is driven by firms that operate in industries that make a more intensive use of relation-specific investments. We find that a one-year increase in trial length decreases the probability to supply customized inputs by 0.4 to 2.1 percentage points in industries at the 25th and 75th percentile of contract intensity, respectively. These effects have an aggregate relevance: for an average level of contract intensity, the hypothetical scenario in which all the courts were as efficient as the best performing one (Vercelli, with an average trial length of 1.03 years) would imply an increase in the aggregate share of subcontracting firms by 2.2 percentage points, from 17.8 to 20%.

In order to improve the causal interpretation of the results, we also exploit a spatial regression discontinuity design (Dell, 2010) that compares the probability to participate to a GVC for firms that are close to the border between two courts that are characterized by different trial lengths. This approach allows us to control for unobservable confounding factors that smoothly vary across space and that might influence the possibility of a firm to participate in GVCs (for example: the different degree of accessibility to transportation facilities or logistical hubs, technological infrastructures installed on the territory, etc.). Results for this exercise are remarkably similar (even quantitatively) to the ones presented in the Rajan-Zingales specification.

Compared with previous studies on the effects of judiciary institutions on international trade (see Nunn and Trefler (2014) for a detailed review of the empirical literature), the advantage of our analysis lies in the fact that we focus on a single country and we exploit the heterogeneity in the quality of institutions within it; this removes all possible concerns related to (nation-wide) omitted variables that correlate with both judiciary efficiency and the participation to GVC (e.g. the availability of certain contractual arrangements or the organization of public administration). In other words, we exploit the heterogeneity of *de facto* local institutions, that are an important determinant for the (often very wide) within-country productivity differentials (Acemoglu and Dell, 2010) and might affect firms along multiple dimensions, such as access to finance, investment and size (Ponticelli and Alencar, 2016).

Italy can be considered as a textbook case of heterogeneous *de facto* local institutions (Cannari, 2009; Cannari and Franco, 2010). When we look at local institutions, North-South Italian divide includes political accountability (Nannicini et al., 2013) and schooling quality (Angrist et al., 2014; Montanaro and Sestito, 2014). Judiciary efficiency is no exception (Giacomelli et al., 2017), despite the fact that local administrations have no role in the organization of local courts and all decisions regarding the allocation of resources are centrally made by an independent body located in Rome. It is important to notice, however, that our empirical analysis will not solely exploit the Italian North-South divide, but also the heterogeneity in the law enforcement within more homogeneous macroareas.

The bulk of the empirical literature on contract incompleteness and international trade uses cross country data. Anderson and Marcouiller (2002) and Berkowitz et al. (2006) show that contract incompleteness can be an important determinate of international trade. Nunn (2007) and Levchenko (2007), further develop this idea and show that countries with better institutional quality have a comparative advantage in the production of goods that are contract intensive. Helpman et al. (2008) estimate a gravity equation to show that countries that share the same legal institutions have a higher probability of establish trade relationships. The use of cross-country data can be problematic because there are two possible sources of institutional quality heterogeneity: countries have different legal system and they differ in institutional enforcement. Our focus on a single country has the advantage of keeping the legal system fixed, while allowing us to focus on the impact of *de facto* institutions within country in the level of law enforcement.

Few other works explore the relationship between the quality of institution and international trade using firm-level data. Using data from 28 developing countries, Ma et al. (2010) show that firms located in areas with better institutional quality export goods that are more contract intensive. Their results replicate the findings of Nunn (2007) using firm level data. Araujo et al.

(2012) show that the importer country's institutional quality affects the export of Belgian firms. With a similar approach, Aeberhardt et al. (2014) use French firm-level data to show that better institutional quality improves the persistence of trade relationships for firms operating in industries with severe contracting problems. Our paper differs from all these studies along two important dimensions: first, they focus on firms' exports and imports rather than firms' purchase or supply of customized goods; second, although from a different perspective, these works assume that institutional quality is country-specific. A notable exception is Feenstra et al. (2013), that exploit cross-provincial variations in contract enforcement effectiveness in China and aggregate trade flows at province level to show that *local* institutions matter for "processing" exports. Our findings complement the rather scarce existing firm-level evidence about the triggers of production fragmentation. In a recent paper, Fort (2017) shows that firms' adoption of ICT technology is an important determinant of the firms' global sourcing strategy.

The reminder of the paper is organized as follows. In section 1.2, we sketch a simple model to provide a theoretical background for our empirical analysis, and in section 1.3 we introduce the Italian legal system. In section 1.4 and 1.5, we describe the data and the empirical strategy. Section 1.6 discusses the results. In section 1.7 we present an alternative strategy that exploits spatial discontinuities to achieve an identification, and we discuss the results. Section 1.8 concludes.

1.2 Theoretical background

In order to analyze the relationship between participation to GVCs and legal enforcement, we consider a simple cash in advance contract (Antràs, 2015) between two firms (a buyer and a seller) over an intermediate input. The economics of the model is quite simple; when the buyer is not able to immediately verify the quality of the good and the contract is not perfectly enforceable, the seller has the incentive to misbehave. This risk implies that the buyer is less willing to sign the contract when law enforcement is weak; this effect is particularly strong if the intermediate input is particularly valuable for the buyer's production.

Model setup

Consider two firms: F (seller) and firm M (buyer). At time t_0, M and F sign the contract; F agrees to sell an intermediate good at price s to be paid at t_0. M has also the opportunity to buy the same good from another supplier F' (in principle, F' can also be a subsidiary of M); however, we assume that if M decides to buy from F', it has to pay a higher price $s' > s$. The fact that the good is paid before the buyer is able to observe the real quality of the supplied good may leave room for seller's misbehavior. We assume, for simplicity, that while F can misbehave, F' cannot.

In period t_1, the production and delivery of the intermediate good take place. If F delivers the exact type of good described in the contract, it pays a production cost $c_1 < s$; however, F may decide to misbehave by delivering a lower quality output, by delaying the delivery, or, in an extreme case, by not delivering the good at all. Faced with this opportunity, F will consider the monetary incentives and the legal implications of such a deviation from the contract. In case of misbehavior, production cost is $c_2 < c_1$ ($c_2 = 0$ if the good is not produced at all). Once the intermediate good is delivered, M immediately incorporates it in the final good. If F behaved according to the contract, M registers a revenue P_1; if misbehavior took place, M's sales are lower (due to either lower quality of the final good or lower total production) $P_2 = P_1 - d$.

The d parameter is crucial in this model; it basically approximates the importance of the intermediate good in the production of the final good. We can assume that the higher d, the more the intermediate good is designed for the buyer's need (in other words, the intermediate good is relationship-specific). From a legal point of view, this generally implies that the larger is d the more complex is the contract signed between M and F. By receiving and incorporating the intermediate good into the final one, at t_1 M understands whether F misbehaved or not.

In the last period (t_2), M may decide to start a lawsuit against F. We assume that M is always able to win and that it will be fully compensated by its loss by receiving $d = P_1 - P_2$. While these hypotheses imply that there is no uncertainty about the final decision of the court, we assume, however, that the cost associated with this lawsuit is *ex-ante* unknown and uniformly distributed between zero and K: $k \sim U[0, K]$. The cost of the lawsuit could be either very low (the court immediately decides in favor of M) or very high (the court takes a long time to make a decision). This cost can be directly linked with trial length. If the court is inefficient and it takes a long time to make a decision, foregone profits could not be immediately reinvested with possible financial losses; moreover, firm M could be forced to divert funds from profitable investments to pay lawyers for a longer time.[1]

[1]In the Italian system, lawyers are generally paid for the time they spend on the case. This implies that, if they have to show up more frequently in court for each case, they get higher payment. It should be noted that it is not possible to fix the cost of the lawyers *ex-ante*, thus arising the uncertainty on the final cost of the lawsuit.

We normalize the litigation cost of F to zero; this basically implies that, on average, litigation cost is higher for M, due to the fact that M is a foreign firm that has to adapt to the legal system of F.

Stage t_2 is extremely streamlined; in fact, we assume that the judicial system is flawless except for trial length. The main reason for this choice of modelling is its simplicity; however, this assumption is not particularly far from truth. As many CEPEJ reports have noted, Italian courts are considered reasonably impartial and independent, but extremely slow.[2]

We assume that firms are risk neutral and all parameters are known by agents; the only uncertain parameter is the cost of the lawsuit k.

Solving the model

The model is solved by backward induction. First note that, if firm F behaves according to the contract, firms' profits are as follows:

$$\pi_M^G = P_1 - s \tag{1.1}$$
$$\pi_F^G = s - c_1 \tag{1.2}$$

where G stands for "good" behavior.

If F deviates, M has to decide on the possible start of a lawsuit. M's profits are:

$$\pi_M^{B,S} = P_1 - s - k \tag{1.3}$$
$$\pi_M^{B,NS} = P_2 - s \tag{1.4}$$

where B stands for a "bad" behavior by F and S (NS) means "sue" ("no sue").

At t_2, M decides to start a lawsuit only if $\pi_M^{B,S} > \pi_M^{B,NS}$, that if only if $P_1 - P_2 = d > k$. Given the uniform distribution of k between zero and K, this occurs with probability $min\{\frac{d}{K}, 1\}$. From now on, we assume that $K > d$ in order to exclude trivial results in which contracts are always enforced.

We now analyze the F's choice to misbehave at t_1. Equation (1.2) shows the payoff when F respects the contract. If it does not, its profits depend on the probability of M's reaction:

$$\pi_F^B = (s - c_2)(1 - \frac{d}{K}) + (s - c_2 - d)\frac{d}{K} = s - c_2 - \frac{d^2}{K} \tag{1.5}$$

[2] See, on this, reports and data available on `http://www.coe.int/T/dghl/cooperation/cepej/default_en.asp`.

The difference between equation (1.5) and (1.2) gives the monetary incentive for F to deviate from the written contract: $c_1 - c_2 - \frac{d^2}{K}$. The incentive is larger when deviation entails a very large saving in terms of production cost $(c_1 - c_2)$; it is also high when the expected cost of the lawsuit for M is high (K). A large compensation to be paid if F is sued (d^2) reduces instead the incentives to deviate for F.[3]

We now abstract from the simple case in which $c_1 - c_2 - \frac{d^2}{K} < 0$ (i.e. there are no incentives for F to deviate) and we assume that a "bad" behavior is always incentive compatible. At t_0, M has to decide whether to sign the contract with the supplier F or the supplier F'. M will sign the contract with F, that is F will access a GVC only if:

$$[P_1 - s - E(k)]\frac{d}{K} + (P_2 - s)(1 - \frac{d}{K}) > P_1 - s' \tag{1.6}$$

where the term at the left of the inequality is the profit for M if the contract with F is signed and the term at the right is the profit if the contract is made with F'. $E(k) = \frac{K}{2}$ is the expected value of the cost of the lawsuit, given its uniform distribution and risk neutrality by firms.

Rearranging (1.6), we obtain the following equation, that describe under which conditions M will access the GVC:

$$s' - s > d(\frac{3}{2} - \frac{d}{K}) = A(K, d) \tag{1.7}$$

The term on the left of the inequality is the cost incentive for M to sign a contract with F. The term on the right is the expected cost due to the risk of deviation, that depends on two parameters: the quality of contract enforcement institutions (K) and the importance of the intermediate input in the production of the final good d. Notice that, for a given difference in the prices of the intermediates $(s' - s)$, the higher $A(K, d)$, the lower the probability for F to access a GVC.[4]

Comparative statics

We are now able to derive two testable predictions of our simple model. The first is the relationship between participation to a GVC and law enforcement. This is done by deriving $A(K, d)$ with respect to K:

---

[3]As we said, if $d \leq K$ the contract is always enforceable.

[4]In principle, M may decide to buy F to avoid the risk of misbehavior. In this case, the term $s' - s$ may also approximate the costs for internalizing production with an acquisition.

$$\frac{\partial A}{\partial K} = \frac{d^2}{K^2} > 0 \tag{1.8}$$

Equation (1.8) is always positive; it basically states that the lower the quality of law enforcement, the more likely it will be that a deviation by F will not be punished by the legal system. As a consequence, the less likely it will be that F participates to a GVC.

The second testable prediction states that when the degree of contract complexity d is higher, it will be less likely for M to sign a contract with a risky supplier in a low quality environment. In formula:

$$\frac{\partial^2 A}{\partial K \partial d} = \frac{2d}{K^2} > 0 \tag{1.9}$$

Equation (1.9) is equivalent to a Rajan-Zingales specification, in which we show that the negative consequences of bad contract enforcement institutions are amplified when d is particularly large.

Discussion on the model

This simple model returns a relationship between law enforcement and participation to GVCs. The model rests on two main assumptions. First, there are no possibilities to repeat the game and acquire reputation, since the model is static. Second, the contract between F and M requires a payment in advance.

As for the first hypothesis, it is true that the possibility to acquire reputation is able to reduce the negative effects of poor formal institutions to participation to GVCs; this would be consistent with the empirical evidence that formal and informal institutions are substitute for economic development (Ahlerup et al., 2009). However, if the population of buyers and sellers is large enough (which is reasonable for tradable goods) and information on agents' behaviors imperfect, this models is basically equivalent to the first stage of a repeated game.[5] However, it should be noted that reputation would just attenuate the estimated effects of the empirical counterparts of equation (1.8) and (1.9); in other words estimated coefficients would represent a lower bound of the actual effect of law enforcement on participation to GVCs.

Regarding the second hypothesis (cash-in-advance contract), if payments may take place after the delivery of the good, the buyer may decide not to

[5]In other words, this can be a reasonable representation of the extensive margin of the participation to a GVC.

pay s if it is not satisfied; in principle, this would make the contract perfectly enforceable. Two things should be noted on this issue. The first is that, as Antràs and Foley (2015) show, most of international transactions generally involve little use of finance and are paid in advance by the buyer. The second is that the economics of the model does not crucially depend on the fact that payment is made in advance but just on the circumstance that the buyer is able to assess the real quality of the intermediate input after assembling the final good and observing its profits. This is not far from real-life transactions, as Antràs (2015) and Midler (2009) point out on a more anecdotal ground.

1.3 The Italian legal framework

We test the two theoretical predictions of the model exploiting the variation in contract enforcement across courts within Italy. This is possible thanks to some institutional features that make the Italian legal system a good empirical setting for our analysis.

First, and most important, in case of litigation the Italian system automatically determines the court in charge of the lawsuit and/or the enforcement of the sentence made by another tribunal. Article 26 of the Italian Code of Civil Procedure states that the court in charge is the one where the majority of the losing firm's properties are located; to draw a parallel with our theoretical model, the court is determined by the location of the supplier of intermediate inputs (F). The court where the firm is located is therefore also in charge of the final enforcement of any decision coming from international disputes, even those issued by foreign courts or international arbitrations. In other words, there is no *choice-of-law* clause, at least for the final enforcement. If the court where the intermediate input suppliers is located is inefficient, the foreign buyer may foresee a substantial reduction of its profits, arising from both the incentive of the supplier to deviate from the contract and the delayed compensation in case of controversy.

Second, in 2011 Italy had 165 tribunal jurisdiction areas, whose boundaries has been set in 1865 after the Italian unification, and that display large differences in contract enforcement.[6] As stated above, World Bank's Doing Business survey shows a very poor performance of the Italian Judiciary system, along with a great variability across different courts. Italy ranks 155th out of 185 countries in terms of trial length: for a first instance decision on a commercial dispute it might take 855 days in Turin and more than 2,000 days in Bari.

[6]The number of courts were sensibly reduced in 2012 as a consequence of large budget cuts to the Judiciary administration.

1.4 Data description

In our empirical exercise, we merge the firm-level data coming from the 9^{th} Census of Industry and Services of 2011 and the information on the efficiency of civil justice provided by the Italian Ministry of Justice. We use a special section of the Census —called "Indagine multiscopo"— which has been administered to the universe of firms with 20 or more employees and to a representative sample of the firms between 3 and 20 employees. Two different sampling schemes have been applied to firms with 3–9 and 10–19 employees: as a result, our sample contains information on the 39% of manufacturing firms with 10 to 19 employees, and on the 19% of those with 3 to 9. Overall we have information on the behavior of 75,006 manufacturing firms.

A notable feature of this dataset is that it exactly identifies the firms that participate to a GVC. Firms were asked whether they supply customized intermediate inputs to foreign companies, that is whether they had relations with other firms under the "*contratto di subfornitura*, L. 192/1998." This type of contract is specific to the case in which the buyer provides design and production criteria to the supplier that performs the physical transformation activities. In what follows, we will also refer to these firms as "international subcontractors". Our data is not perfect, though. The most relevant limitation is that we do not observe the content of the contract, the value of the transaction and the identity of the partners. We are only able to identify, among firms, those that supply customized inputs abroad. A similar variable was used by Fort (2017) in her analysis on the sourcing strategies of US firms. When the firm belongs to a foreign group, we are also able to know if the majority of these supplies were realized on behalf of other firms of the group; in that case, the firm is dropped from the sample, since our contract incompleteness argument does not apply within the boundaries of a group. Our database contains 14,983 firms supplying customized goods to foreign firms outside their group. These pieces of information are integrated by other data on the firm's location, sector of activity, number of employees, revenues and value added.

As already explained, we proxy the quality of contract enforcement with trial length. This is calculated by using the caseflow data provided by the Ministry of Justice.[7] For each of the 165 Italian judicial districts, we compute:

$$D_t = \frac{P_t + P_{t+1}}{E_t + F_t} \tag{1.10}$$

[7]This is also the official measure used by the National Statistical Agency (ISTAT) to provide national figures, and have been already used by Giacomelli and Menon (2016), among others.

where P are pending cases at the beginning of the year, F are the new cases filed throughout the year, and E are the cases ended with a judicial decision or withdrawn by the parties during the year. The data only refer to ordinary civil proceedings and are expressed in years. In order to get rid of possible idiosyncratic volatility in the series of trial length, we take the average of the index between 2002 and 2007 as our measure of institutional quality.

As shown by equation (1.9) we expect the effect of local courts' efficiency on the probability of accessing GVC to scale up with the contract intensity of the goods provided. This is approximated by a sectoral index of relationship specificity derived from the Rauch's classification (Rauch, 1999). We measure the contract intensity as the share of differentiated products produced within each sector, using both the liberal and the conservative classification.

Table 1.1 contains some basic statistics on our sample of firms; besides the statistics on the full sample, we also report those on the universe of manufacturing firms with 20 or more employees. In our sample, the average size in terms of employees is 32.6, while average revenues and value added amount to roughly 9 and 2 million Euros; average value added per worker amounts to 44.6 thousand Euros. 56% of the firms in our sample is an exporter; a slightly lower percentage of firms supply customized inputs, but only 17.8% does it abroad. It has to be noted that most of the firms operate in contract-intensive sectors, as captured by an average Rauch index above 80% under both definitions. Bigger firms also display a higher labor productivity; moreover, they also have a greater probability of exporting and of supplying customized inputs, either domestically or abroad. The average contract intensity for the full sample and the big firms alone are remarkably similar. Interestingly, the average trial length to which firms are exposed does not significantly differ across firm size, suggesting that bigger firms do not tend to sort in the jurisdiction of more efficient courts.

In table 1.2, we show how firm characteristics vary with the exporting status. In line with the theory, domestic firms are on average smaller and less productive, while exporters are characterized by the largest average size (both in terms of employees and revenues) and by the highest value added per worker; this relative rank —although with greater magnitudes— is preserved in the more selected sample of big firms. In a similar fashion, firms that supply customized inputs abroad are larger and more productive than firms that do it only domestically, though being comparable to firms that do not sell customized goods at all in most dimensions. Most interestingly for our purposes, domestic producers are on average located in judicial districts characterized by a higher length of civil proceedings; this pattern also holds for international suppliers, located in better courts than domestic suppliers and firms not involved in subcontracting, hinting at a potential role played by local court efficiency on the exporting behaviour of a firm.

To further explore this point, Figure 1.3 displays two maps highlighting the geographical distribution of firms that supply customized inputs and the duration of civil proceedings by judicial district. The comparison of the two panels suggests a relevant negative correlation between the two variables. Firms participation into GVC is more frequent in the North and the Centre, and is limited to very narrow zones of the South, which is instead characterised by a longer duration of civil trials. The same correlation can be explored in a regression framework: table 1.3 displays the results of a court-level regression of the share of international subcontractors on the average duration of civil trials. The correlation is negative, meaning that less efficient courts are associated with a lower share of international suppliers of customized inputs, and remains significant even when controlling for a North-South dummy; significance vanishes when controlling for region fixed effects, though the coefficient remains negative even under this more demanding specification.

1.5 Empirical strategy

To test the first prediction of the theoretical model, that is if the effectiveness of contract enforcement affects the probability for a firm to engage in international subcontracting, we estimate the following equation:

$$y_{ic} = \alpha + \beta TL_c + \gamma X_i + \varepsilon_{ic} \tag{1.11}$$

where y_{ic} is a dummy indicating whether firm i, located in the jurisdiction of court c, supply customized inputs abroad. TL_c is a measure of quality of law enforcement (trial length, measured in years) in court c: the higher TL_c, the more inefficient the court. Finally, X_i is a vector containing firm level controls: it always includes size class fixed effects, that aim at controlling for the different sampling schemes used to collect data (remember that our data encompass the universe of the firms with 20 employees or more, while it is a representative sample of the firms with 3–9 and 10–19 employees). Besides this, X_i always includes a dummy for the firms located in the Southern regions; the purpose of adding this control is to ensure that our estimates are not simply picking up the effect of the North-South divide in both economic development and court efficiency.[8] Depending on the specification, X_i may also contain additional controls, such as industry fixed effects (4 digit of the

[8]A civil trial lasts on average 3.3 years in the South, while it lasts 2 years in the rest of Italy.

Ateco2007 classification), log employment, value added per worker, and a dummy that indicates whether the firm belongs to a group. As discussed in previous sections, we expect the estimate of β to be negative if a higher level of court inefficiency —as measured by the average trial length— reduces a firm's probability of entering in an international GVC as a subcontractor.

To check the second theoretical prediction, however, we have to bring our empirical analysis a step further, exploring the existence of an heterogeneous effect of institutional quality across different levels of contract intensity. To do that, we estimate the following equation:

$$y_{ic} = \alpha + \beta_1 TL_c + \beta_2 CI_i + \beta_3 TL_c \times CI_i + \gamma X_i + \varepsilon_{ic} \tag{1.12}$$

Here CI_i measures the degree of contract intensity of the sector in which firm i operates, according to the classification proposed by Rauch (1999). Besides the already mentioned controls, under this specification the matrix X_i may also include a set of court dummies. The inclusion of these controls is intended to control for potential confounders that may be common to all firms belonging to the same court; nonetheless, adding court dummies will cause the effect of trial length *per se* (β_1) to be absorbed. In the same way, since the Rauch contract intensity index varies at the industry level, the inclusion of industry (4 digit) fixed effects will absorb the β_2 coefficient.

In the most requiring specification, we will therefore only be capturing the effect of judicial quality mediated through the level of contract intensity of the sector in which the firm operates, as captured by coefficient β_3. Under this approach, the access to GVC is explained by the interaction between an industry and an area characteristic. This specification closely follows the empirical strategy used by Rajan and Zingales (1998) to test the relationship between financial development and dependence on external financing, or the one used by Nunn (2007) in his study on law enforcement as a source of country level comparative advantages. In our case, we expect $\beta_3 < 0$ if the negative effect of court inefficiency (that is, a longer trial length) on the probability of a firm to supply customized inputs abroad scales up with the contract-intensity of the activities conducted, holding fixed all other characteristics.

A few things should be noted in the estimation of equations (1.11) and (1.12). Under this empirical strategy, we do not have a source of exogenous variation in the data: this implies that the causal interpretation of the coefficients rests on the discussion of potential sources of biases in our estimates.

First of all, there could be omitted variables which may influence the access to GVC, while being at the same time correlated with trial length. For example, contract intensive sectors generally display a higher average productivity (for example, because firms in those sectors make a more intensive use of skilled labor), which is a crucial predictor for the access to international markets and which, in Italy, tends to characterize the areas where the law enforcement is more efficient (North). Mechanisms of this kind may create a downward bias in our estimates. For this reasons, we control for firm-level determinants of international subcontracting like size and productivity. Moreover, the literature highlights additional factors that explain the probability of accessing a GVC, and that might in principle be influenced by the quality of local institutions, such as the degree of electronic codifiability of the activities performed by the firms, or the availability/endowment of ICT infrastructures (Fort, 2017). Though these mechanisms might be at work in Italy, they are not likely to affect our estimates: on one hand, product codifiability mostly varies across sectors, and we control for industry fixed effects at a very fine level (4 digits); on the other hand, differential endowments of ICT technologies at the local level are controlled for by the geographic fixed effects included (in the most demanding specification, we include court-level fixed effects). In general, the issue of omitted variable bias is more precisely tackled in section 1.7, by using a spatial regression discontinuity approach to control for smoothly-varying factors between two neighboring courts.

The second issue relates to reverse causality. Areas in which international sourcing is very diffuse may successfully lobby the Italian Ministry of Justice to maintain a good contracting environment by keeping there the most efficient judges, court officers and clerks, thus negatively affecting our coefficients of interest. While we cannot exclude this occurrence, this issue looks much more relevant in cross-country analyses rather than in within-country (and, especially within-Italy) regressions. In Italy, decisions on the composition of local courts are made by the High Council of the Judiciary (HCJ), a central governing body of the judicial system whose independence is guarantee by the Italian Constitution (articles 105 to 107).[9] HJC decides according to the dispositions of two major laws: the first is the Royal Decree n. 12 issued on January 30th, 1941, the second is Law 195, published on March 24th, 1958. Both laws were issued in a completely different economic setting, well before problems related to international production fragmentation could even arise. This said, HCJ still retains some discretionary powers in the assignment of judges; yet, HCJ is an extremely independent body, whose autonomy is warranted by the Constitution and jealously defended by its components.

Finally, a third issue is related to the problems of sorting or self-selection.

[9]Two-thirds of the HJC are made by judged, which are elected by all Italian judges. One-third is instead elected by the Parliament among University Professors of Law or Lawyers.

Firms that supply customized inputs abroad might be induced to relocate in areas in which courts are more efficient. This would generate a negative bias in our coefficients of interest. We can test this issue by checking whether areas with a better quality of law enforcement tend to be more specialized in contract-intensive sectors. Figure 1.4 plots the average contract intensity at court level against the average trial length. We use both liberal and conservative definition of contract intensity as defined by Rauch (1999). Although this evidence is not conclusive about the role of sorting, the correlation between the two measures is weak, thus suggesting that sorting is unlikely to be the main driver of our results.[10]

1.6 Results

The estimation of equation (1.11) yields the baseline results reported in Table 1.4. As discussed above, the coefficient of interest is the one attached to trial length, which is our measure of institutional quality. As expected, the coefficient on trial length is negative throughout all the specifications, remaining highly significant as geographic and firm level controls are added. The magnitude of the coefficient remains remarkably stable, as we add industry fixed effects and other firm-level controls such as size (employees) and productivity (value added per worker). The estimate in column (3) tells us that a one-year increase in the length of civil trials would reduce the probability of supplying customized inputs abroad by 1.9 percentage points.

The inclusion of a dummy for the firms located in the Southern regions ensures that the estimated effect is not biased by secular differences between North and South, that influence both the quality of judicial institutions and the propensity of firms to participate in GVCs; the coefficient on this dummy tells us that being located in the South is associated with a 5% lower probability of supplying customized inputs abroad.[11] As expected, size and productivity positively correlate with the probability of engaging in international subcontracting. Being part of a group significantly increases the probability of entering a GVC.[12]

[10]Giacomelli and Menon (2016) also provide evidence on the lack of sorting in response to court efficiency in Italy.

[11]All results are confirmed when we exclude Southern regions, that is when we eliminate a set of geographical units characterized by both weak contract enforcement and lagging economic conditions; results also hold when we introduce a finer set of regional (NUTS2) dummies.

[12]Results are robust even to the inclusion of dummies identifying exporters and subcontractors (both domestic and international), with the aim of controlling for self-selection into either status. Note that these controls may in turn be plausibly affected by trial length

We then explore the possibility that the effects of institutional quality on the participation to GVCs may be heterogeneous, depending on the contract intensiveness of the industry in which the firm operates, as suggested by the theoretical model. To test that, we then turn to the estimation of equation 1.12; the results are displayed in Table 1.5. Column (1) reports the estimates for the most parsimonious specification, which includes trial length, contract intensity and their interaction, along with the dummy for South and a set of sector fixed effects (2 digits level).[13] As expected, the coefficient on Rauch's measure is positive, confirming that the probability of entering a GVC is increasing in the scope for differentiation of the goods produced. The coefficient on trial length, instead, is non-significant and very close to zero. The effect of institutional quality is mediated by the extent of contract intensity in the sector where a firm operates, as suggested by the coefficient on the interaction term, which is negative and significant: low contract enforcement has a detrimental effect on the participation to GVCs, increasingly so for industries characterized by a higher contract intensity. As more geographic and firm controls are added, the coefficient on trial length stays close to zero, while the one on the interaction term remains negative and significant throughout.[14]

In columns (2)–(4) we progressively add the firm controls, and the industry- and court- level fixed effects, which absorb the coefficient on contract intensity and on trial length, respectively. As more controls are added, the magnitude of the coefficient on the interaction term slightly reduces, though it does not suffer a huge drop: according to the most demanding specification in column (4), a one-year increase in trial length reduces the probability to operate as an international subcontractor by 0.4 percentage points for firms belonging to sectors at the 25th percentile of the (liberal) Rauch classification; the fall amount to 2.1 percentage points for industries at the 75th percentile.[15] A similar magnitude can be obtained when we look at the conservative definition of the Rauch index. To quantitatively assess the aggregate relevance of efficient legal institutions on firms' participation to GVCs, we calculate how would the overall share of firms engaged in international

(i.e. they could be outcome variables themselves), and are therefore very likely to configure as bad controls, thus leading to a downward bias of our parameter of interest. Despite decreasing in magnitude as expected, the estimate remains negative and significant even under this more requiring (and likely wrong) specification.

[13]Contract intensity varies at 4 digit level.

[14]Similar results are obtained if we use the conservative version of Rauch classification in place of the liberal one.

[15]The distribution of the Rauch indexes is characterized by a substantial mass of sectors producing only differentiated products; as a consequence, the 75th percentile of the distribution of the Rauch index (both liberal and conservative) is equal to 1.

subcontracting vary, in the hypothetical situation in which all courts were as efficient as the best performing one. To do that, we equalize the trial length across tribunals, setting it to the one of the Vercelli court, which displays the lowest average trial duration in Italy (1.03 years). We then calculate for each tribunal the implied effect on GVC participation, keeping the contract intensity fixed at the average value registered at the court level. Finally, we aggregate these effects, appropriately weighting them for the number of firms in each court, to obtain a back of the envelope quantification of the aggregate effect: if all the courts were as efficient as the Vercelli one is, the share of subcontracting firms in Italy would raise by 2.2 percentage points, from 17.8 to 20%.

In Table 1.6 we check the robustness of our estimates, addressing two different concerns. First, the dependent variable in equations 1.11 and 1.12 equals to one if the firm is an international subcontractor; zero is instead attributed to both non-subcontractors and to subcontractors that only operate domestically. A possible concern relates to the fact that the negative coefficient we found may actually depend on the self-selection into the subcontractor status rather than on the effect of ineffective contract enforcement on the probability to enter a GVC. In order to reject this hypothesis, we re-run our estimates on subcontractors only (columns (1) and (2)). The estimates are robust to this sample cut, remaining negative and significant, with point estimates that are even higher than those presented before.

A second conspicuous concern relates to the presence of multi-plant firms. In section 1.5 we have discussed that in that case the court in charge of the execution may vary according to the relative size of the firm's assets across plants. The Census collects information on the number of plants belonging to each firm, and this allows us to exclude multi-plant firms from our analysis, in order to control for the potential identification problems related to the multi-localization of assets. Results are presented in the last two columns of table 1.6. Column (3) reports the estimates of equation (1.12) without interaction, while column (4) adds heterogeneous effects. The results are in line with a potential attenuation bias in our baseline estimates, due to measurement error: restricting our sample to single plants only, the point estimates are in fact slightly larger in modulus than the baseline ones.

Finally, we make a sanity check on the effects of local institutions on the exporter status as well (this is equivalent to a firm level estimate of Nunn (2007)). Table 1.7 replicates the previous specifications, using the dummy for exporters as a dependent variable. Results are similar both in sign and significance, but much stronger in magnitude. This consideration applies both to the specification with trial length alone and to the one with the interaction. In the latter case, a one-year increase in trial length would reduce the probability of being an exporter by 0.7 percentage points for firms at the 25th percentile of the (liberal) Rauch classification, and by 3.9 percentage points for those at the 75th percentile.[16]

1.7 Exploiting spatial discontinuities

In order to improve the causal interpretation of our results, we adopt an alternative identification strategy that exploits the fact that the quality of the institutions discontinuously varies at the border dividing one court from another. The basic idea behind this estimation framework is that —if we consider two firms at two different sides of a court's boundary— they will only differ in the average trial length they are facing, once other geographically-varying confounding factors are controlled for; this empirical setting will therefore allow us to assess the causal impact of being subject to a less efficient court on the firm probability of engaging in international subcontracting. It is worth remembering that the geographical boundaries of Italian courts do not coincide with those of other administrative divisions such as the provinces, so that there is no risk of picking up the effect of other institutional discontinuities.[17]

Building on other empirical works that exploit geographic discontinuities (see, for example, Holmes (1998); Black (1999); Dell (2010); Gibbons et al. (2013)), we adopt an estimation framework that allows us to identify an effect across multiple borders. We partition the boundaries of each court in several segments, one for every neighboring court;[18] henceforth, when we talk about borders we will be referring to these segments. Ideally, we would like to observe a firm's localization to determine its distance from the nearest border; since we are not able to observe a firm's exact position, we proxy it with the centroid of the municipality in which the firm is located. This plausibly leads to a measurement error, that should —if anything— attenuate our estimates.

For each border, it is straightforward to rank the two neighbouring tribunals according to their average trial length. As a consequence, for each border we can identify the set of municipalities (hence, of firms) located in

[16] A possible concern is that these estimates are driven by the international subcontractors, whose negative relationship with trial length has already been documented; however, dropping these firms leaves the results on exporters basically unchanged.

[17] As discussed in section 1.3, the boundaries of the 165 tribunal jurisdictions were established in 1865 after the Italian unification, long before the creation of the provinces as self-standing administrative organizations.

[18] As a consequence, we get as many segments as the pairs of neighbouring tribunals we are able to single out in the administrative map of the courts.

the tribunal with the (relatively) better or worst institutions.[19] If we define belonging to a (relatively) bad court as our treatment of interest, we see that the probability for a firm of being treated experiences a discrete unitary jump at the border between tribunals. The linear distance from the nearest border will therefore be the running variable in our empirical exercise. We conventionally attach negative values of distance to municipalities (and firms) lying on the bad side of a border, and positive values to those belonging to the good side: moving from positive distance values towards the border, a firm will thus get treated when the distance drops down to zero. To test the continuity of the running variable across the borders, we implemented the test put forth by McCrary (2008), whose failure would be hinting at possible manipulations of the treatment assignment; the test is strongly rejected, thus providing another indirect evidence of no sorting across the border.

Having structured the problem like this, we will be estimating the following regression in the first place:

$$y_{icb} = \alpha + \gamma bad_{cb} + f(\text{border distance}_{ib}) + \delta_b + \delta_c + \beta X_i + \varepsilon_{icb} \qquad (1.13)$$

where y_{icb} is a dummy indicating whether firm i, located in court c near border b has engaged in international subcontracting; bad_{cb} is the treatment dummy, equal to 1 if court c's performance is poorer than the one of the neighboring court belonging to the same border b. $f(\text{border distance}_{ib})$ is a polynomial of the distance from the nearest border, which aims at controlling for all the factors smoothly varying as a function of distance even within a single court. We explore different functional forms for this polynomial, by progressively increasing its order; moreover, in order to allow for the maximum degree of flexibility of our estimates, we allow the parameters of the polynomial to vary across treated and non-treated units. δ_b and δ_c are borders and courts fixed effects; as a consequence of pooling together multiple borders, it is important to include these fixed effects, since our empirical setup is only valid within border and controlling for relative court efficiency.

Though in principle our identification strategy based on geographic discontinuities does not require additional controls, we still include them in certain specifications to improve the precision of our estimates, by providing additional balancing for firms across the border. X_i is the matrix containing firm level controls: as in equation 1.11, it includes 4 digits industry fixed effects and size class fixed effects.[20] Moreover, in certain specifications X_i

[19] We are not able to find a neighbouring court for every municipality. This is the case, for example, of municipalities lying on the national boundaries or on some segments of the coastline. We are forced to exclude these municipalities from our analyses, which implies dropping around 10,000 firms.

[20] Industry fixed effects capture mostly technological and market-related characteristics

may also contain additional controls, such as log employment, value added per worker and a dummy for firms belonging to a group.[21]

The estimate of coefficient γ in equation 1.13 will yield the average effect of being subject to a slower tribunal on a firm's probability of engaging in international subcontracting. Since we have previously argued that the quality of judicial institutions is more binding for firms operating in a contract-intensive sector, we then verify whether the estimated average effect displays some heterogeneity across different levels of contract intensity. Based on recent contributions on the estimate of heterogeneous effects in a regression discontinuity design (Becker et al., 2013; Accetturo et al., 2014), we estimate the following equation:

$$\begin{aligned} y_{icb} =& \alpha + \gamma_1 bad_{cb} + \gamma_2 CI_i + \gamma_3 bad_{cb} \times CI_i + \\ &+ f(\text{border distance}_{ib}) + \delta_b + \delta_c + \beta X_i + \varepsilon_{icb} \end{aligned} \tag{1.14}$$

where CI_i —again based on Rauch (1999) classification— proxies for the contract intensity of the sector in which firm i operates. For our argument to hold, the contract intensity should amplify the negative effect of belonging to a less efficient tribunal, hence, we would expect coefficient γ_3 to be negative.

Table 1.8 displays the results obtained estimating equation (1.13) across different specifications. Columns (1)-(3) report the estimates of the most parsimonious specifications, where we do not add any firm-level control and we just let the order of the distance polynomial vary. Using a linear distance polynomial yields, as expected, a negative coefficient, though not significantly different from zero. The linear distance polynomial is, however, a pretty rough approximation, and when we use a quadratic polynomial —as in column (2)— the coefficient on treatment increases in modulus and becomes significant; the estimates remain basically unchanged when we use a third order polynomial. Since we know that the results obtained under high-order polynomials are likely to be misleading (see Gelman and Imbens (2014)), we take the the quadratic distance polynomial specification as our preferred one, and in column (4) and (5) we add some firm level control. Results do not change appreciably and significances increase. The estimates

that determine the role played by the sector in the international fragmentation of production (as for the codifiability of operations discussed above). Size class fixed effects control for the different sampling schemes used to collect data, which vary according to the firm's dimension (3–9, 10–19 and 20+ employees).

[21] As in section 1.6, we have also tried including a dummy for exporters and one for subcontractors, that we have already argued to be bad controls in our framework. When we add them to the model, the point estimates decrease as expected, though remaining negative and significant.

attached to firm-level controls remain basically unchanged with respect to those presented in Table 1.4.

To provide a graphical illustration of our exercise and to show that the results are not driven by outliers, Figure 1.5 replicates the results on collapsed data. More specifically, we first net our dependent variable from border, court, industry and size class fixed effects; then we partition the border distance in equally spaced bins. The scatter plot in Figure 1.5 displays the bin averages of our netted dependent variable, along with a second order polynomial fit on distance. The fit suffers a sharp downward shift at the border: the difference the two lines returns the estimated effect of being subject to a lower-quality judiciary environment (corresponding to negative distance values).

It is difficult to compare the results obtained under this framework with those discussed in the previous section; this is because in our geographical discontinuity design the treatment is discrete and captures relative differences in court efficiency across borders, while in the previous case we had a direct quantification of the effect of trial length on subcontracting. We can, however, try to provide a rough quantification of the estimates obtained with spatial discontinuities. The estimate displayed in column (5) tells us that, on average, a firm located in a court characterized by a lower institutional quality has a 2.1% lower probability of engaging in international subcontracting. This coefficient has to be interpreted in the light of an average trial length difference across borders of about 0.6 years. Hence, if we wanted to force a comparison with the results presented in Table 1.4, we would state that the effect turns out to be larger when we identify it under the geographic discontinuity framework.[22]

We then turn to estimating equation (1.14), which takes into account the contract intensity of the sector in which the firms operate. Results are displayed in Table 1.9, and the specifications adopted are the same as those discussed in the previous table. The coefficient attached on treatment remains non-significant throughout all the specifications; the sign is negative, once we adopt a non-linear distance polynomial. As in Table 1.5, the coefficient on contract intensity *per se* is positive and significant in the specifications where it is not absorbed by sector fixed effects. What we are primarily interested in, however, is the coefficient on the interaction term, which is negative, significant and satisfactorily stable across all the specifications: a firm located in the jurisdiction of an inefficient court has a lower probability of supply-

[22]Remember that in Table 1.4 we had found that a one-year increase in trial length would make the probability of engaging in international subcontracting decrease by 1.9 percentage points.

ing customized inputs abroad, and this effect is increasingly negative for the firms operating in more contract-intensive sectors. Figure 1.6 again provides a graphical illustration of our results: the exercise is equivalent to the one presented above, but now we split the sample between firms in sectors with low and high contract intensity (below the 25^{th} or above the 75^{th} percentile of the Rauch (1999) index). Panel (a) shows that at the border little or no effect can be identified for firms with low contract intensity; for those at the top of the Rauch classification, instead, the effect emerges clearly, as the fitted line shifts downward in the negative region. On top of this, it is interesting to notice that —irrespective of their position around the border— firms operating in more contract-intensive sectors on average enjoy a higher probability of engaging in international subcontracting.

As for the magnitude of the effects, we again provide a rough comparison with the results obtained in the previous section. The coefficient on the interaction term in column (5) means that the treatment effect scales up with the level of contract intensity, reducing the probability of entering GVCs by 0.3 percentage points for firms at the 25th percentile of Rauch distribution and by 1.7 for firms at the 75th percentile. This compares again to an average cross-border difference in trial length by 0.6 years. With a simple back of the envelope calculation, we could state that if the treatment amounted to a one-year difference in trial length, the effect on the probability of international subcontracting would be negative by 0.5 and 2.8 percentage points for firms at the 25th and 75th percentile of the Rauch distribution, respectively. These magnitudes are slightly larger, but completely in line with what we had shown in Table 1.5.

In order to assess the soundness of these results, we conduct some robustness checks based on the sample of firms included in our regressions and on the specification of the distance polynomial. A typical robustness check when dealing with regression discontinuity designs involves replicating the results on progressively smaller samples, by shrinking the geographical buffer considered around the border; this check is intended to rule out the possibility of picking up spurious effects that are driven by distant observations. In our setup, we expect this robustness check to even strengthen our estimates. This is because in our empirical setup we have assumed the relevant border for comparison to be the nearest one from each municipality. However, this might not always be the case: in the first place, we have considered the linear distance from the border, while we have disregarded other physical and infrastructural factors (e.g. mountains or roads) that might make a border less accessible than it is through the simple fly distance criterion; second, the assignment of a municipality to a border is increasingly arbitrary as the municipality is closer to the geographical center of the tribunal jurisdiction area.

As we shrink the buffer around the borders, these concerns become less and less relevant, as the arbitrariness of the border-municipality match vanishes; as a consequence, measurement error should also drop, so that we might expect our estimates to also gain in precision. We therefore replicate the specification in column (5) of Table 1.9 over progressively smaller samples, identified by reducing the buffer around each border. Results are displayed in Table 1.10. As a matter of fact, the estimates of both the treatment effect and the interaction term remain negative throughout the subsamples. The coefficient on the interaction gains in significance as the buffer around the border shrinks, and the point estimates even rise in magnitude when considering a 5km buffer.

Finally, we replicate the results in Table 1.9 using a bi-dimensional spatial polynomial in both latitude and longitude, in the spirit of Dell (2010). This is a more demanding version of the distance polynomial used so far, and better controls for all the factors that vary smoothly across space.[23] Results are displayed in Table 1.11. The estimates of the treatment effect *per se* almost drop to zero and remain non significant as before. Nonetheless, the coefficients on the interaction term remain strikingly similar —both in terms of point estimates and of significance— to the ones obtained using the simpler distance polynomial.

1.8 Concluding remarks

In this paper we study if and to what extent the quality of judiciary institutions at the local level influence the probability for a firm to enter GVCs. To do that, we use a special section the 2011 Italian census, targeted at providing statistical information on Italian firms' participation to the international fragmentation of production; this data source has the merit of exactly identifying the firms that sell customized inputs abroad (international subcontractors), that are the main subject of our study.

We exploit the fact that Italy displays a substantial amount of heterogeneity in the quality of local institutions. We focus in particular on judiciary efficiency —measured by the length of civil trials at the court level— as the relevant dimension that is likely to reverberate on the capacity of a firm to sign customized inputs provisioning contracts with foreign counterparts. We argue that the inherent riskiness of a contract gets inflated when the judicial system is unable to guarantee an efficient and timely enforcement of the rule of law. Hence, firms subject to the jurisdiction of an inefficient court may

[23] To make an example, a second order polynomial in latitude and longitude would be defined as $lat + lon + lat \times lon + lat^2 + lon^2$.

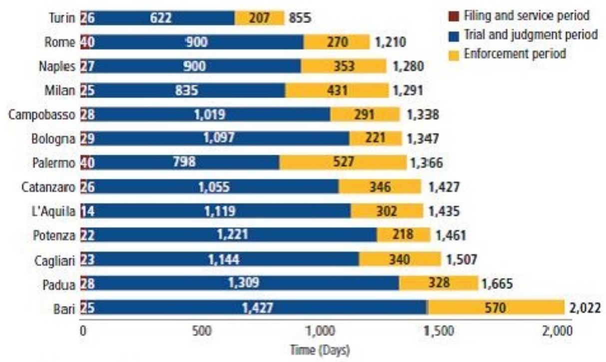
face additional difficulties in entering GVCs as a subcontractor. This is a particularly relevant issue for a country like Italy, which ranks very low in cross-country comparisons on the effectiveness of contract enforcement.

Our results show that firms located in courts with a higher trial length in civil disputes are less likely to supply customized inputs to foreign firms: a one-year increase in trial length reduces the probability of entering a GVC by 1.9 percentage points. The effect is stronger for firms operating in sectors characterized by a strong contractual activity (which typically are the sectors that require high relationship-specific investments). For the firms at the 25th percentile of the distribution of the contract-intensity index, the effect of a one-year increase in trial length is almost negligible (0.4 percentage points), while it is much larger for those lying at the 75th percentile (2.1 percentage points). For an average level of contract intensity, the hypothetical scenario in which all the courts were as efficient as the best performing one (Vercelli, with an average trial length of 1.03 years) would imply an increase in the aggregate share of subcontracting firms by 2.2 percentage points, from 17.8 to 20%.

To corroborate our results, we adopt a more demanding identification strategy that exploits the fact that the quality of the institutions discontinuously varies at the border dividing one court from another. Under this empirical framework, we compare the firms lying at the two sides of each court boundary, and identify the treatment effect of being located in the side characterized by the worst quality of judicial institutions. The results confirm the negative and significant effect of trial length on the probability of supplying customized inputs abroad; also in this case, the effect is stronger for firms operating in contract-intensive sectors. As for the magnitude of these effects, a rough quantification exercise suggests that they are comparable across specifications. The empirical setup based on the spatial discontinuity across tribunals returns slightly larger estimates: one-year difference in trial length would negatively affect the probability of international subcontracting by 0.5 and 2.8 percentage points for firms at the 25th and 75th percentile of the contract intensity distribution, respectively.

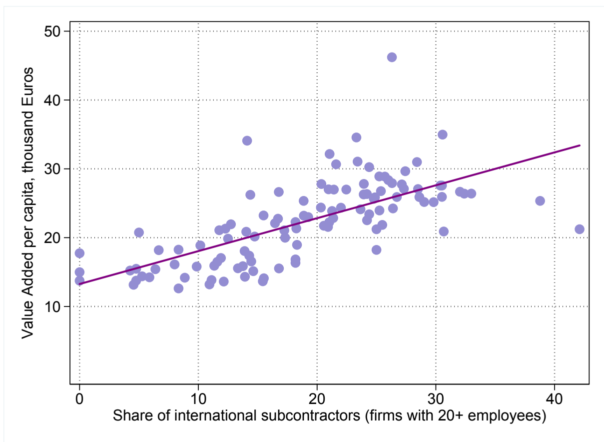
Figures and Tables

Figure 1.1: Heterogeneity in the efficiency of Italian civil justice



Source: Doing Business report.

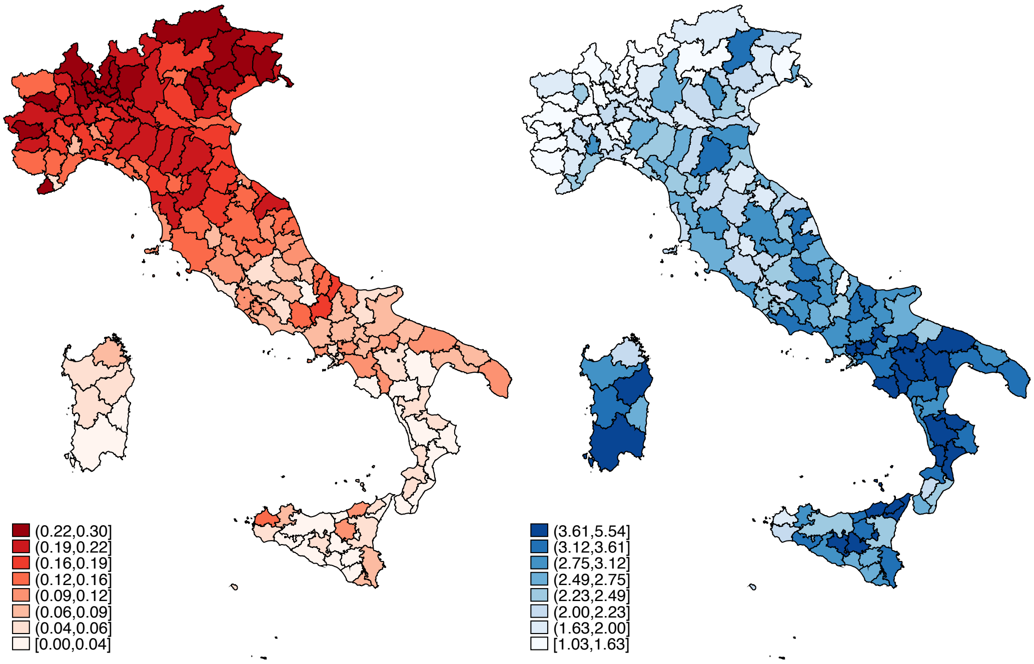
Figure 1.2: Participation to GVCs and Value Added per capita at provincial level in Italy



Source: 9th Census of Industry and Services and ISTAT National Accounts.

Figure 1.3: The geography of subcontracting and judiciary efficiency

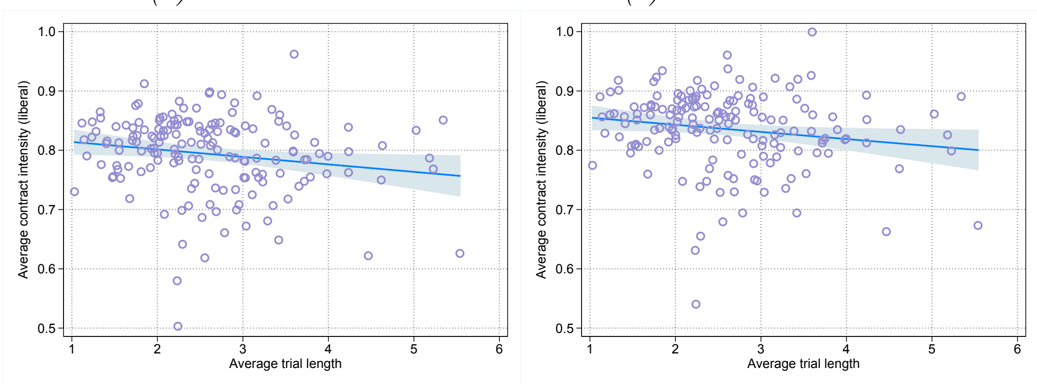
(a) Share of subcontractors *(b) Average trial length*



Notes: 9*th* Census of Industry and Services and Italian Ministry of Justice. The left panel shows the share of firms participating in GVCs within each court.

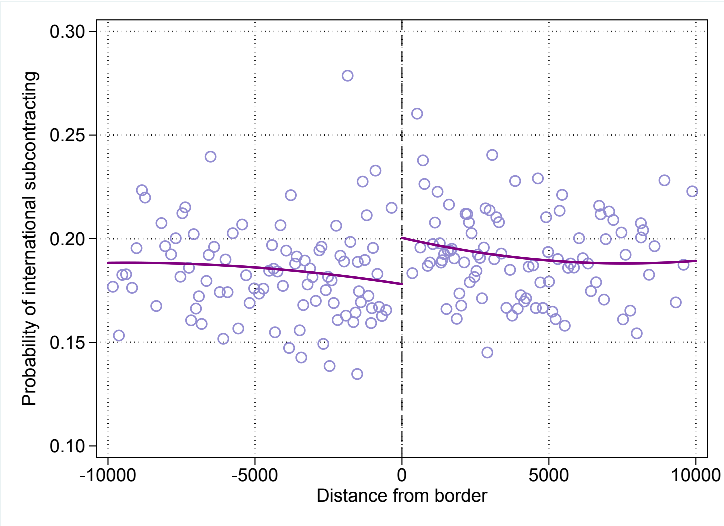
Figure 1.4: Controlling for sorting

(a) Rauch -liberal *(b) Rauch - conservative*



Source: 9*th* Census of Industry and Services and Italian Ministry of Justice. *Notes:* Each dot represents a court. On the y-axis we report the average contract intensity (according to Rauch (1999) liberal classification) across firms within a court. The x-axis represents the average trial length at the court level, expressed in years.

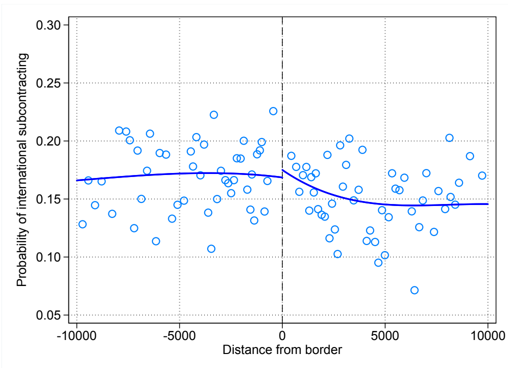
Figure 1.5: Spatial discontinuity, treatment effect at the border



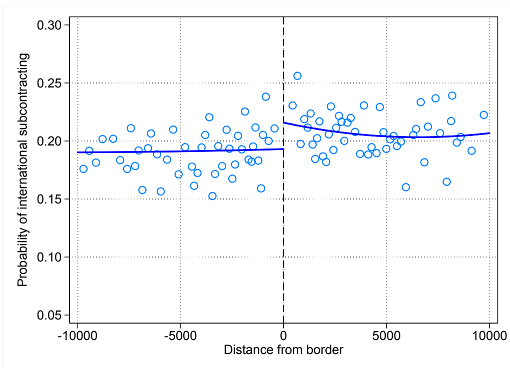
Notes: The distance around the border has been partitioned in 200 equally-spaced bins. The dots represent the average within each bin of the dependent variable (dummy for international subcontracting), netted of border, court, industry and size class fixed effects. The line is a 3rd order polynomial fit.

Figure 1.6: Spatial discontinuity, heterogeneous effect by contract intensity

(a) Below 25^{th} *percentile*



(b) Above 75^{th} *percentile*



Notes: In each panel, the distance around the border has been partitioned in 100 equally-spaced bins. The dots represent the average within each bin of the dependent variable (dummy for international subcontracting), netted of border, court, industry and size class fixed effects. The line is a 3rd order polynomial fit.

Table 1.1: Summary statistics

	Full sample		20+ employees	
	Mean	Std. dev.	Mean	Std. dev.
Dummy exporter	0.561	0.496	0.679	0.467
Dummy international subcontractor	0.178	0.383	0.236	0.425
Dummy subcontractor	0.527	0.499	0.568	0.495
Employees	32.621	154.454	68.339	239.728
Revenues (million €)	9.015	108.867	20.141	171.465
VA (million €)	1.977	11.956	4.389	18.631
VA per worker (thousand €)	44.600	75.794	52.999	38.806
Trial length (in years)	2.312	0.721	2.275	0.696
% Rauch liberal	0.820	0.289	0.822	0.287
% Rauch conservative	0.861	0.272	0.864	0.269

Source: 9*th* Census of Industry and Services and Italian Ministry of Justice.

Table 1.2: Summary statistics by firm status

	#	empl	revs (M€)	VA (M€)	VApw (th€)	trial length
			(a) Full sample			
Domestic only	32,257	18.43	3.47	0.86	36.26	2.44
Exporting	41,208	43.73	13.36	2.86	51.13	2.21
No subcontracting	34,744	32.82	10.14	2.08	43.34	2.35
Domestic subcontractor	25,273	27.93	6.76	1.51	41.88	2.33
International subcontractor	2,573	46.52	12.02	3.13	54.97	2.17
Both	10,875	39.58	9.97	2.47	52.52	2.17
			(b) 20+ employees			
Domestic only	9,422	43.86	9.17	2.23	42.78	2.42
Exporting	19,970	79.89	25.32	5.41	57.82	2.21
No subcontracting	12,704	75.30	24.95	5.11	54.33	2.29
Domestic subcontractor	9,622	58.73	15.51	3.36	47.71	2.33
International subcontractor	1,439	75.21	19.44	5.17	59.63	2.17
Both	5,627	67.29	17.38	4.33	57.35	2.18

Notes: 9*th* Census of Industry and Services and Italian Ministry of Justice. The first column shown the numerosity of each group, while all the other statistics are group averages. The last column displays the average trial length in the judicial districts where the firms are located.

Table 1.3: Regression at court level

	(1)	(2)	(3)
Trial length	-0.0496***	-0.0201***	-0.0084
	[0.0051]	[0.0058]	[0.0054]
Dummy South		-0.0852***	
		[0.0099]	
Region FE	N	N	Y
R^2	0.383	0.584	0.770
Obs.	165	165	165

Notes: Variables are court-level averages. Robust standard errors in brackets. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.4: Baseline regressions

	(1)	(2)	(3)
Trial length	-0.0215***	-0.0190***	-0.0191***
	[0.0063]	[0.0046]	[0.0046]
Dummy South	-0.0850***	-0.0580***	-0.0532***
	[0.0090]	[0.0070]	[0.0068]
Log employees			0.0322***
			[0.0039]
VA per worker			0.0001**
			[0.0001]
Business group			0.1332***
			[0.0082]
Size class dummy	Y	Y	Y
Industry FE	N	Y	Y
R^2	0.031	0.069	0.077
Obs.	73,465	73,465	73,435

Notes: The dependent variable is a dummy for firms exporting customized goods. Business group is a dummy for firms belonging to a group. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. Standard errors in brackets clustered at the court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.5: Heterogeneous effects by contract intensity

	(1)	(2)	(3)	(4)
Trial length	0.0033	0.0004	0.0005	
	[0.0056]	[0.0056]	[0.0056]	
Contract intensity	0.0661***			
	[0.0151]			
TL $\times$ CI	-0.0296***	-0.0236***	-0.0239***	-0.0210***
	[0.0062]	[0.0062]	[0.0062]	[0.0062]
Dummy South	-0.0630***	-0.0583***	-0.0535***	
	[0.0072]	[0.0069]	[0.0068]	
Log employees			0.0322***	0.0318***
			[0.0039]	[0.0039]
VA per worker			0.0001**	0.0001**
			[0.0001]	[0.0001]
Business group			0.1331***	0.1322***
			[0.0082]	[0.0081]
Size class dummy	Y	Y	Y	Y
Sector FE	Y	N	N	N
Industry FE	N	Y	Y	Y
Court FE	N	N	N	Y
R^2	0.053	0.069	0.078	0.083
Obs.	73,465	73,465	73,435	73,435

Notes: The dependent variable is a dummy for firms exporting customized goods. The Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Business group is a dummy for firms belonging to a group. Sector fixed effects defined according to the Ateco2007 classification at 2 digits; industry fixed effects are instead defined at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. Standard errors in brackets clustered at the court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.6: Robustness checks: restricting sample to subcontractors and single plants

	Only subcontractors		Only single plants	
	(1)	(2)	(3)	(4)
Trial length	-0.0357***		-0.0203***	
	[0.0082]		[0.0049]	
TL $\times$ CI		-0.0334***		-0.0239***
		[0.0125]		[0.0071]
Log employees	0.0933***	0.0939***	0.0509***	0.0505***
	[0.0056]	[0.0055]	[0.0043]	[0.0043]
VA per worker	0.0002	0.0002	0.0001*	0.0001*
	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Business group	-0.0672***	-0.0691***	0.1270***	0.1258***
	[0.0089]	[0.0090]	[0.0111]	[0.0110]
Dummy South	-0.0846***		-0.0536***	-0.1490***
	[0.0114]		[0.0072]	[0.0066]
Size class dummy	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Court FE	N	Y	N	Y
R^2	0.132	0.144	0.081	0.088
Obs.	38,707	38,707	57,501	57,501

Notes: The dependent variable is a dummy for firms exporting customized goods. In the interaction term, the Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Business group is a dummy for firms belonging to a group. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. Standard errors in brackets clustered at the court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.7: Export as a dependent variable

	(1)	(2)	(3)	(4)
Trial length	-0.0398***	-0.0386***	0.0054	
	[0.0099]	[0.0098]	[0.0114]	
Contract Intensity			0.0875***	
			[0.0272]	
TL × CI			-0.0590***	-0.0390***
			[0.0119]	[0.0094]
Dummy South	-0.1405***	-0.1300***	-0.1427***	
	[0.0151]	[0.0149]	[0.0165]	
Log employees		0.1247***	0.1286***	0.1241***
		[0.0045]	[0.0047]	[0.0045]
VA per worker		0.0003	0.0003	0.0003
		[0.0002]	[0.0002]	[0.0002]
Size class dummy	Y	Y	Y	Y
Sector FE	N	N	Y	N
Industry FE	Y	Y	N	Y
Court FE	N	N	N	Y
R^2	0.204	0.223	0.157	0.238
Obs.	73,465	73,435	73,435	73,435

Notes: The dependent variable is a dummy for exporters. The Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Sector fixed effects defined according to the Ateco2007 classification at 2 digits; industry fixed effects are instead defined at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. Standard errors in brackets clustered at the court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.8: Spatial regression discontinuity, baseline results

	(1)	(2)	(3)	(4)	(5)
Treatment	-0.0050	-0.0224**	-0.0224*	-0.0223***	-0.0217***
	[0.0116]	[0.0101]	[0.0124]	[0.0077]	[0.0080]
Log employees					0.0308***
					[0.0042]
VA per worker					0.0000**
					[0.0000]
Business group					0.1356***
					[0.0081]
Border FE	Y	Y	Y	Y	Y
Court FE	Y	Y	Y	Y	Y
Size class FE	N	N	N	Y	Y
Industry FE	N	N	N	Y	Y
Spatial polynomial	1st	2nd	3rd	2nd	2nd
R^2	0.026	0.026	0.026	0.076	0.084
Obs.	63,283	63,283	63,283	63,282	63,252

Notes: The dependent variable is a dummy for firms exporting customized goods. Treated firms are those located on the side of the border with a longer trial length. Business group is a dummy for firms belonging to a group. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. One-dimensional spatial polynomial based on the distance from the border; the order of the polynomial is reported at the bottom of the table. Standard errors in brackets clustered at the border and court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.9: Spatial regression discontinuity, heterogeneous effects by contract intensity

	(1)	(2)	(3)	(4)	(5)
Treatment	0.0119	-0.0059	-0.0062	-0.0081	-0.0082
	[0.0145]	[0.0136]	[0.0150]	[0.0110]	[0.0116]
Contract Intens.	0.0770***	0.0770***	0.0771***		
	[0.0089]	[0.0088]	[0.0089]		
Treatment × CI	-0.0210**	-0.0209**	-0.0210**	-0.0174*	-0.0166*
	[0.0104]	[0.0104]	[0.0104]	[0.0097]	[0.0098]
Log employees					0.0308***
					[0.0042]
VA per worker					0.0000**
					[0.0000]
Business group					0.1356***
					[0.0081]
Border FE	Y	Y	Y	Y	Y
Court FE	Y	Y	Y	Y	Y
Size class FE	N	N	N	Y	Y
Industry FE	N	N	N	Y	Y
Spatial polyn.	1st	2nd	3rd	2nd	2nd
R^2	0.028	0.028	0.028	0.076	0.084
Obs.	63,283	63,283	63,283	63,282	63,252

Notes: The dependent variable is a dummy for firms exporting customized goods. Treated firms are those located on the side of the border with a longer trial length. The Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Business group is a dummy for firms belonging to a group. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. One-dimensional spatial polynomial based on the distance from the border; the order of the polynomial is reported at the bottom of the table. Standard errors in brackets clustered at the border and court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.10: SRD robustness check: restricting sample around the border

	(1) full sample	(2) 20 km	(3) 15 km	(4) 10 km	(5) 5 km
Treatment	-0.0082	-0.0066	-0.0082	-0.0134	-0.0061
	[0.0116]	[0.0110]	[0.0112]	[0.0150]	[0.0219]
Treatment × CI	-0.0166*	-0.0175*	-0.0159*	-0.0172*	-0.0249**
	[0.0098]	[0.0094]	[0.0087]	[0.0087]	[0.0117]
Log employees	0.0308***	0.0313***	0.0307***	0.0314***	0.0352***
	[0.0042]	[0.0043]	[0.0043]	[0.0048]	[0.0053]
VA per worker	0.0000**	0.0000**	0.0000**	0.0000**	0.0000*
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Business group	0.1356***	0.1380***	0.1382***	0.1416***	0.1377***
	[0.0081]	[0.0080]	[0.0082]	[0.0092]	[0.0146]
Size class FE	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Border FE	Y	Y	Y	Y	Y
Court FE	Y	Y	Y	Y	Y
Spatial polyn.	2nd	2nd	2nd	2nd	2nd
R^2	0.084	0.084	0.085	0.085	0.092
Obs.	63,252	62,339	59,776	52,093	30,740

Notes: The dependent variable is a dummy for firms exporting customized goods. Treated firms are those located on the side of the border with a longer trial length. The Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Business group is a dummy for firms belonging to a group. Sample is progressively restricted, selecting firms belonging to an increasingly narrower buffer around the court border. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. One-dimensional spatial polynomial based on the distance from the border; the order of the polynomial is reported at the bottom of the table. Standard errors in brackets clustered at the border and court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 1.11: SRD robustness check: two-dimensional spatial polynomial

	(1)	(2)	(3)	(4)	(5)
Treatment	0.0082	0.0087	0.0083	0.0065	0.0058
	[0.0117]	[0.0117]	[0.0118]	[0.0104]	[0.0109]
Contract Intens.	0.0768***	0.0766***	0.0764***		
	[0.0087]	[0.0087]	[0.0087]		
Treatment × CI	-0.0210**	-0.0209**	-0.0205*	-0.0175*	-0.0166*
	[0.0103]	[0.0103]	[0.0104]	[0.0096]	[0.0098]
Log employees					0.0308***
					[0.0042]
VA per worker					0.0000**
					[0.0000]
Business group					0.1354***
					[0.0082]
Border FE	Y	Y	Y	Y	Y
Court FE	Y	Y	Y	Y	Y
Size class FE	N	N	N	Y	Y
Industry FE	N	N	N	Y	Y
Spatial polyn.	1st	2nd	3rd	2nd	2nd
R^2	0.028	0.028	0.028	0.076	0.084
Obs.	63,283	63,283	63,283	63,282	63,252

Notes: The dependent variable is a dummy for firms exporting customized goods. Treated firms are those located on the side of the border with a longer trial length. The Rauch index in its liberal version is used as a measure of contract intensity; results do not vary when the conservative version is used. Business group is a dummy for firms belonging to a group. Industry fixed effects defined according to the Ateco2007 classification at 4 digits. Three size classes are defined, reflecting the different sampling schemes adopted by the "Indagine multiscopo" survey: 3–9 employees; 10–19 employees; 20+ employees. Two-dimensional spatial polynomial based on latitude and longitude; the order of the polynomial is reported at the bottom of the table. Standard errors in brackets clustered at the border and court level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

References

Accetturo, A., G. de Blasio, and L. Ricci (2014). A Tale of an Unwanted Outcome: Transfers and Local Endowments of Trust and Cooperation. *Journal of Economic Behavior & Organization 102*(C), 74–89.

Acemoglu, D. and M. Dell (2010). Productivity Differences Between and Within Countries. *American Economic Journal: Macroeconomics 2*(1), 169–188.

Aeberhardt, R., I. Buono, and H. Fadinger (2014). Learning, Incomplete Contracts and Export Dynamics: Theory and Evidence from French Firms. *European Economic Review 68*, 219–249.

Ahlerup, P., O. Olsson, and D. Yanagizawa (2009). Social Capital vs Institutions in the Growth Process. *European Journal of Political Economy 25*(1), 1–14.

Anderson, J. E. and D. Marcouiller (2002). Insecurity and the Pattern of Trade: An Empirical Investigation. *The Review of Economic and Statistics 84*(2), 342–352.

Angrist, J. D., E. Battistin, and D. Vuri (2014). In a Small Moment: Class Size and Moral Hazard in the Mezzogiorno. NBER Working Paper Series n. 20173.

Antràs, P. (2003). Firms, Contracts, and Trade Structure. *The Quarterly Journal of Economics 118*(4), 1375–1418.

Antràs, P. (2015). *Global Production: Firms, Contracts, and Trade Structure.* Princeton University Press.

Antràs, P. and C. F. Foley (2015). Poultry in Motion: a Study of International Trade Finance Practices. *Journal of Political Economy 123*(4), 853–901.

Antràs, P. and E. Helpman (2004). Global Sourcing. *Journal of Political Economy 112*(3), 552–580.

Araujo, L., G. Mion, and E. Ornelas (2012). Institutions and Export Dynamics. CEPR discussion paper.

Becker, S. O., P. H. Egger, and M. von Ehrlich (2013). Absorptive Capacity and the Growth and Investment Effects of Regional Transfers: A Regression Discontinuity Design with Heterogeneous Treatment Effects. *American Economic Journal: Economic Policy 5*(4), 29–77.

Berkowitz, D., J. Moenius, and K. Pistor (2006). Trade, Law, and Product Complexity. *The Review of Economics and Statistics 88*(2), 363–373.

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics 114*(2), 577–599.

Cannari, L. (2009). The Mezzogiorno and Regional Policies. *Bank of Italy Workshop and Conferences 2*, 7–744.

Cannari, L. and D. Franco (2010). The Mezzogiorno and Italian Economic Policy. *Bank of Italy Workshop and Conferences 4*, 1–215.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*(6), 1863–1903.

Feenstra, R. C., C. Hong, H. Ma, and B. J. Spencer (2013). Contractual versus Non-Contractual Trade: The Role of Institutions in China. *Journal of Economic Behavior & Organization 94*, 281–294.

Fort, T. C. (2017). Technology and Production Fragmentation: Domestic versus Foreign Sourcing. *The Review of Economic Studies 84*(2), 650–687.

Gelman, A. and G. Imbens (2014). Why High-Order Polynomials should not be used in Regression Discontinuity Designs. NBER Working Paper n. 20405.

Giacomelli, S. and C. Menon (2016). Does Weak Contract Enforcement Affect Firm Size? Evidence from the Neighbour's Court. *Journal of Economic Geography lbw030*, 1–32.

Giacomelli, S., S. Mocetti, G. Palumbo, and G. Roma (2017). La giustizia civile in Italia: le recenti evoluzioni. Forthcoming in Bank of Italy Occasional Papers.

Gibbons, S., S. Machin, and O. Silva (2013). Valuing School Quality using Boundary Discontinuities. *Journal of Urban Economics 75*, 15–28.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics 123*(2), 441–487.

Holmes, T. J. (1998). The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders. *Journal of Political Economy 106*(4), 667–705.

Lazerson, M. (1999). The Firms that Feed Industrial Districts: A Return to the Italian Source. *Industrial and Corporate Change 8*(2), 235–266.

Levchenko, A. A. (2007). Institutional Quality and International Trade. *Review of Economic Studies 74*(3), 791–819.

Ma, Y., B. Qu, and Y. Zhang (2010). Judicial Quality, Contract Intensity and Trade: Firm-level Evidence from Developing and Transition Countries. *Journal of Comparative Economics 38*(2), 146–159.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142*(2), 698–714.

Midler, P. (2009). *Poorly Made in China.* John Wiley & Sons.

Montanaro, P. and P. Sestito (2014). The Quality of Italian Education: a Comparison between the International and the National Assessments. *Bank of Italy Occasional Papers 218*, 4–42.

Nannicini, T., A. Stella, G. Tabellini, and U. Troiano (2013). Social Capital and Political Accountability. *American Economic Journal: Economic Policy 5*(2), 222–250.

Nunn, N. (2007). Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade. *The Quarterly Journal of Economics 122*(2), 569–600.

Nunn, N. and D. Trefler (2014). Domestic Institutions as a Source of Comparative Advantage. In *Handbook of International Economics*, Volume 4, Chapter 5, pp. 263. Elsevier.

Ponticelli, J. and L. S. Alencar (2016). Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil. *The Quarterly Journal of Economics 131*(3), 1365–1413.

Rajan, R. G. and L. Zingales (1998). Financial Dependence and Growth. *American Economic Review 88*(3), 559–586.

Rauch, J. E. (1999). Networks versus Markets in International Trade. *Journal of International Economics 48*(1), 7–35.

UNCTAD (2013). Global Value Chains and Development. Available at: http://unctad.org/en/PublicationsLibrary/diae2013d1_en.pdf.

Chapter 2

Allocative Efficiency and Finance

Joint with Andrea Linarello and Enrico Sette, Bank of Italy.

2.1 Introduction

Productivity is the engine of economic growth. After the Great Recession, which has been triggered by a credit crunch in many developed countries, a growing body of economic research studied to what extent credit shocks affect aggregate productivity. Negative credit shocks can impact aggregate productivity through several channels. First, they can lower firm-level productivity, as they exacerbate credit constraints, preventing firms from investing, hiring workers and innovating. Second, credit shocks could increase firm exit, which may benefit aggregate productivity, to the extent that low productivity firms are forced to leave the market. Third, negative credit supply shocks affect the entry rate of firms: typically, the productivity of entrants is higher during downturns (Lee and Mukoyama, 2015), but negative credit shocks could attenuate this positive selection, and may delay the growth of new entrants (Midrigan and Xu (2014)). These channels, however, do not account for the full impact of finance on aggregate productivity: if credit constraints force low productivity firms to shrink, unconstrained high productivity firms may be able to expand, thus fostering the reallocation of production factors towards more productive uses.

In this paper we measure the effect of credit supply shocks on aggregate productivity. Importantly, we go beyond the study of the impact of credit supply shocks on firm-level productivity alone, but we also study its effect through the reallocation of labor across firms, and through the exit and entry margins. We are in an ideal position to address this question, since we have access to a unique dataset including the universe of Italian manufacturing firms covering the period 2003-2014 (BdI-ISTAT). This is crucial to obtain a complete picture of the reallocation process and of the entry and exit of firms. Throughout the paper, our empirical approach will be guided by the Melitz and Polanec (2015) decomposition of aggregate productivity. This allows us to measure the effect of credit supply shocks on productivity through different channels: i) the impact of the credit shock on the growth of incumbent firms' productivity; ii) the contribution of individual firms to the covariance between market share and productivity (which measures the extent of reallocation); iii) the extensive margin, looking at the impact of the credit shock on entry and exit and on the productivity of entrants and exiters relative to the incumbent firms. In the period under analysis, the negative contribution deriving from the massive fall in firm-level productivity has been mitigated by the exit of less productive firms, and especially by the reallocation component.

We isolate credit supply shocks at the industry-province level using detailed microdata from the Italian Credit Register using the procedure proposed in Amiti and Weinstein (2013). In a nutshell, we regress the growth rate of credit by each bank in each sector-province controlling for a full set of sector-industry-time and bank-time fixed effects. The latter represent the credit supply shocks, which we then aggregate at the sector-province level, using the share of credit of each bank in each sector-province. This approach allows us to purge our estimates from demand effects, which typically affect the dynamics of credit (Khwaja and Mian, 2008; Amiti and Weinstein, 2013), as the credit supply shocks are, by construction, orthogonal to the firms' demand for credit.

Importantly, our data encompass both a period in which the Italian economy experienced good economic growth and the two deep recessions following the default of Lehman (2009–2010) and the European sovereign debt crisis (2012–2014). This allows us to study the impact of credit supply shocks on productivity during financially-driven recessions, and to test for differential effects of credit shocks in good as opposed to crises times. Moreover, it gives us the chance to roughly quantify the overall impact of the credit crunch on aggregate productivity, distinguishing the various channels through which its effects unfolded.

Our findings show that a restriction in credit supply affects aggregate productivity growth through various channels. On one hand, it depresses productivity growth at the firm level. On the other hand, it provides a positive and sizable contribution through the reallocation component. This happens because, as a consequence of a credit restriction, less productive firms shrink in size, thus losing employment shares in favour of more productive ones. Finally, a weakening of credit supply growth has significant but modest effects in terms of the net demography margin: on the one hand, it increases the positive contribution of exit to aggregate productivity, mainly by reducing the productivity of exiting firms relative to the incumbents and by increasing the exit rate; on the other hand, it further lowers the negative contribution of entry, because entrants in a period of worse credit supply availability are on average less productive relative to the incumbents. Overall, the positive contribution of reallocation and exit has more than compensated the negative effect that credit crunch has had on within-firm productivity and entry: if after 2008 credit supply had continued to grow at its pre-crisis rates, the growth rate of aggregate productivity would have been lower by about 2.5 percentage points per year in the period 2008–2014. This has come at the cost of the expulsion of a substantial fraction of the workforce, that has triggered a reallocation of employment shares towards more productive activities; our estimates show that roughly 100,000 jobs (the 12% of the observed reduction in manufacturing employment) were destroyed among incumbent firms, as a consequence of the credit restriction.

Our findings adds to the large literature on misallocation and productivity. Following the pioneering contribution of Hsieh and Klenow (2009), who find sizable misallocation of inputs in China and India, a large literature has sought to identify the reasons behind the observed frictions in credit markets, in law enforcement, or in the allocation of production factors across firms, and to quantify their relative importance in shaping the pattern of TFP growth. Financial frictions in particular have been the focus of a large and growing literature. Buera and Shin (2013) find that financial frictions have a large impact along the transition to the steady state, prolonging the adverse consequences of the initial resource misallocation. In addition Moll (2014) suggests that financial frictions amplify TFP shocks in the short run, and firms find it difficult to save out of borrowing constraints. Larrain and Stumpner (2012) find that a capital account liberalization decreases resources misallocation by improving the allocation of finance. Midrigan and Xu (2014) challenge these findings suggesting that financial frictions play a limited role in the misallocation of resources, and they do so by creating a distortion in entry and exit rates. A recent work by Gopinath et al. (2017) finds that following the beginning of the European monetary union, the decline in the real interest rate, often attributed to the euro convergence process, lead to a significant decline in sectoral total factor productivity, as capital inflows are misallocated toward firms that have higher net worth but are not necessarily more productive.This effect has been especially pronounced in Spain. Two recent work focusing in Italy study the effect of credit supply on TFP.

Manaresi and Pierri (2016) show that an expansion in the credit supply increases both input accumulation and firms' ability to generate value added for a given level of inputs, in this way enhancing productivity. More indirectly, Schivardi et al. (2017) find evidence of zombie lending in Italy during the financial and sovereign debt crises, although they find limited real effects of this kind of credit misallocation: sales, investment and employment of non-zombie firms are hardly affected by the intensity of zombie lending.

Our work contributes to this literature in two ways. First, we explore the effect of credit market frictions on various components of aggregate productivity, thus shedding light on the channels (average firm productivity, reallocation, entry/exit margin) through which credit shocks affect productivity. Our finding of an increase of reallocation in correspondence with a credit restriction is also consistent with the view that major restructuring episodes tend to be concentrated in recessions, as in Schivardi (2003). Second, we use a unique dataset covering the universe of Italian firms, which allows us to measure the extent of reallocation and selection (entry/exit margins) along the entire firm size distribution. In order to give a comprehensive assessment of the effect of credit supply shocks on aggregate productivity, it is essential to observe the smaller and less productive firms, since they are likely to be most immediately and severely affected by harsher market conditions. Our findings suggest that negative shocks to bank credit on one hand depress average productivity at the firm level, but on the other hand contribute to "cleanse" the economy through the reallocation of resources from low to high productivity firms (Foster et al., 2016), thus dampening the drop in aggregate productivity growth observed during recessions. It has to be stressed, however, that such a process is not necessarily welfare-improving, especially in the short term. The observed positive contribution of reallocation arises from the expulsion of a significant amount of workers —especially from smaller and less productive firms— that is not immediately re-absorbed by bigger firms in manufacturing. What primarily drives our results is therefore a reallocation of employment shares rather than a reallocation of workers.

The paper is organized as follows. Section 2.2 presents the data used in this paper, illustrates the estimation method of the credit supply shocks, and shows some basic stylized facts on firm data and the estimated shocks. Section 2.3 documents the dynamics of aggregate labor productivity and presents the results of the Melitz and Polanec (2015) decomposition, providing some suggestive evidence on the connection between the conditions of credit supply and the extent of reallocation and selection. In section 2.4 we use firm-level data to analyze the effect of a credit supply shock on the components of aggregate productivity, shedding some light on the underlying mechanisms. Section 2.5 roughly quantifies the overall impact of the 2008–2014 credit restriction on aggregate productivity. Section 2.6 concludes.

2.2 Data

The paper relies on two different data sources. The first is a firm-level dataset that covers the universe of manufacturing firms that were active for at least 6 months in a given business year from 2003 to 2014. The construction of the dataset is the result of a joint collaboration between the Bank of Italy and the Italian National Statistical Agency (ISTAT); it combines the information of the Italian Register of Active Firms (ASIA) with data retrieved from statistical, administrative and fiscal sources. The dataset contains information on firms' location, incorporation date, industry classification (Nace rev. 2), number of employees and sales.[1] Data on value added are only available from 2005. We deflate the data on sales and value added to 2010 prices, using sector-level price indexes for sales and value added, respectively. In the spirit of Geurts and Van Biesebroeck (2014), we exploit administrative information to obtain a measure of entry and exit of firms purged from errors related to ID changes, spin-offs and mergers.

The quality of this data can be gauged by comparing them with National Accounts data. Panel (a) of Figure 2.1 compares the value of production from National Accounts with the total value of sales from ASIA dataset, both evaluated at current prices.[2] The two series are very similar over the entire period of observation. The National account series usually remains above the ASIA data, because the former includes estimates of the underground economy and illegal workforce; occasionally, the National Account series lies below the ASIA one, as a consequence of the dynamics of inventories, that are not accounted for by our dataset. The similarity with the National Accounts also emerges when looking at the growth rates, as shown in panel (b); the two series are remarkably close in the central part of our sample and in correspondence to the great trade collapse episode.

The second data source we use is the comprehensive Italian Credit Register, a database owned by the Bank of Italy, which contains data on all individual bank-borrower relationships with an exposure of at least 75,000 Euros until 2008, and 30,000 since 2009. The Credit Register lists outstanding balances of loan amounts at the lender-borrower level aggregated into 3 categories: overdraft loans, term loans, loans backed by receivables, and

[1]See Abbate et al. (2017) for a detailed description of the dataset.

[2]The comparison is made at current prices in order to exclude the discrepancies deriving from the use of price deflators at different levels of disaggregation.

it also flags non-performing loans. Banks routinely use the Credit Register to assess the creditworthiness of current and prospective borrowers, which ensures a high quality of the data. Unique identifiers of banks and borrowers allow us to track them over time. The Credit Register contains both granted (committed) credit and actually used (drawn) credit. We focus on the former as it represents a better measure of credit supply, while the latter is heavily influenced by borrowers' decisions to utilize available credit. We select loans to non-financial firms to compute credit granted by each bank at the industry (Nace rev. 2) and province level. Provinces are local administrative units of a size comparable to that of U.S. counties, which are the relevant market for deposits and small business lending, according to the Italian antitrust authority.

2.2.1 Stylized facts on firm demography and performance

During our sample period (2003-2014) the manufacturing sector was interested by a secular process of structural transformation and shrunk significantly.[3] Table 2.1 reports descriptive statistics on the firms in our sample. The number of firms steadily declined: in 2014 there were about 100,000 firms less than in 2003. As a consequence, the number of employees dropped by roughly 850,000 units. Average firm size —measured in terms of employees per firm— experienced an increase, almost exclusively concentrated in the first half of our sample. The financial crisis heavily contributed to depress the economic performance of Italian manufacturing firms, whose sales started suffering sizable swings.

Aggregate labor productivity, measured as real sales per worker, strongly decreased during the global financial (2007–09) and the sovereign debt (2012–13) crises. The double-dip recession had a severe effect on Italian aggregate labor productivity, which in 2014 was only slightly above its 2007 levels.[4]

2.2.2 The credit supply shock: estimation and basic facts

We use the Credit Register data to estimate bank-specific credit shocks, applying the methodology of Greenstone et al. (2014). We aggregate credit

[3]In the same period, the service sector expanded, instead.

[4]An extended version of this dataset, that we do not use in this paper due to data incompleteness, shows that aggregate labor productivity had been falling —though less intensely— during a previous episode of economic downturn occurred in the years 2002–03. Elaborations on these years are available upon request.

granted by each bank at the province-sector time level and we estimate the following model:

$$\Delta \ln(L_{bpst}) = \alpha_{bt} + \gamma_{spt} + \epsilon_{bspt} \tag{2.1}$$

where $\Delta \ln(L_{bpst})$ is the log change in credit granted by bank b to sector-province sp at time t. α_{bt} are a set of bank $\times$ time fixed effects and γ_{spt} are a set of sector-province $\times$ time fixed effects. Model 2.1 compares the growth of credit from different banks lending to the same sector-province in any year. The sector-province $\times$ time fixed effects control for changes in demand and economic conditions at the sector-province level in each year, while the bank $\times$ time fixed effects α_{bt} are the components of the credit dynamics that are common to each bank b across the credit relationships observed, and can therefore be interpreted as bank-specific credit supply shocks.[5] The set of bank-time fixed effects, α_{bt}, identifies a supply-induced change in credit under the assumption that at the sector-province-time level there is no bank-specific demand for credit, so that the set of sector-province-time fixed effects fully control for changes in demand and in the riskiness and economic prospects of the sector-province. Under this condition, these shocks are uncorrelated with with any characteristics of the firms and of the markets in which the banks operate. In general, as we work at the 2-digit Nace sector and provinces are local administrative units comparable to U.S. counties, this assumption is not particularly restrictive. It could be violated if a bank specialized in financing a certain industry in a given province. Even in this case, though, the set of bank $\times$ time effects can still be interpreted as a supply-side shock (Amiti and Weinstein (2013)).

We work at the sector-province level rather than at the firm-level because our sample includes the universe of firms, while firm-level bank shocks would be available only to firms which appear in the Credit Register, i.e. those which had outstanding loans above 75,000 euros (30,000 since 2009). However, this is not necessarily a drawback for our purposes. Since we are primarily interested in investigating across-firms reallocation rather than within-firm productivity, working at the sector-province level allows us to gain insights on the effects that unfold through the firm size distribution, besides the obvious advantage of taking into account the universe of active firms.

We then aggregate these bank-specific shocks to obtain a measure of the evolution of credit supply at the sector and province level. More specifically,

[5]This approach to identify the bank-lending channel at the firm-level has been first proposed by Khwaja and Mian (2008).

we compute our credit supply shock as:

$$CSS_{spt} = \begin{cases} \sum_b \theta^{sp}_{b,1999}\widehat{\alpha}_{bt}, & \text{if } t \leq 2007 \\ \sum_b \theta^{sp}_{b,2006}\widehat{\alpha}_{bt}, & \text{if } t > 2007 \end{cases} \tag{2.2}$$

where θ^{sp}_{bt} is the market share of bank b in sector s and province p in year t. These shares are computed aggregating the loans in the Credit Register at the sector-province level, as in the computation of the growth rates.

This is a weighted average of the bank $\times$ time fixed effects, in which weights are the share of credit of each bank at the sector-province level as of 1999 and 2006. Due to the relatively long time span covered by our data, we have chosen to let the weights vary to obtain a cleaner measure of the bank shocks as of before the financial and the sovereign debt crises. On the one hand, fixing the market shares at their 1999 levels would make the estimated credit supply shock progressively less informative on the actual propensity to lend, as years move away from 1999; on the other hand, letting the weights vary every year would make our credit supply measure potentially endogenous to the economic performance within each sector $\times$ province cell.[6] Moreover, this formulation of the supply shock comes particularly handy when we split the sample in the two periods before and during the financial crisis: the last year of each subsample is equally distant from the year in which the weights are set.

Since the bank shocks α_{bt} are identified up to a constant scaling factor, the credit supply shock cannot be attached an absolute quantitative interpretation. The differences among banks supply shocks both cross-sectionally and over time are, instead, preserved. For the sake of clarity, suppose we have a sector-province cell for which we estimate a credit supply shock of 5 and -5 at time t and $t + 1$, respectively: we are not able to state whether credit supply actually expanded or shrunk in the two periods (since it is not possible to derive the reference level), but we can assert that the growth rate of credit supply decreased by 10 percentage points; the same comparison can be performed across sector-province cells. This means that —if we were interested in investigating the elasticity of a certain variable to the dynamics of credit supply in a regression framework— it would be perfectly fine to use our estimated credit supply shock as an explanatory variable, since the unknown reference level would not affect the estimate of the elasticity, and would instead be absorbed by the constant.

[6]We have checked the robustness of the estimates presented in section 2.4 by using a credit supply shock obtained both by fixing weights as of 1999 and by letting weights vary across years. Results are basically unchanged. In the former case, the magnitude of the estimated coefficients is slightly attenuated, while in the latter it is slightly inflated.

Table 2.3 shows basic descriptive statistics of the credit supply shocks obtained as shown in equation 2.2. It is apparent that after the outbreak of the global financial crisis in 2008 the propensity of financial intermediaries to lend dramatically declined, with even greater intensity in the years of the sovereign debt crisis.[7]

The dispersion of the credit supply shock across sectors and provinces slightly increased after the crisis. The distribution of the bank shocks by sectors, shown in Figure 2.2, suggests that the drop in credit supply growth during the crisis has been stronger in food, machinery, plastic and metal industries. Differences across sector, however, are very limited, with the credit supply shock being bounded between 4 and 5% before the crisis and between -7 and -6 % after its outbreak. This stylized fact goes in favor of our argument of the estimated credit supply shock being uncorrelated with sector-specific characteristics. This argument is further corroborated by the fact that, at the firm level, Bofondi et al. (2017) found little differences in the banks' lending policies during the sovereign debt crisis across certain firm characteristics such as size, level of indebtedness and capacity to repay interests.

Figure 2.3 shows the distribution of the credit supply shock across provinces. Importantly, most of the provinces which experienced the most negative shocks housed banks which later turned into troubled banks (e.g. Marche, Florence and Arezzo, several provinces in Veneto). The concentration of the most negative shocks in the Center and in the North, the areas that stood the crisis relatively better, suggests that the negative shocks are mostly driven by the strength of bank balance sheet operating in the province, rather than by province characteristics.

To provide further support to the identification of the bank-shocks, we test their correlation with key bank balance-sheet characteristics which are regarded as major drivers of banks' propensity to lend. To this aim, we exploit balance sheet information from the Supervisory Reports submitted by banks to the Bank of Italy. Results, shown in table 2.2 indicate that banks with lower interbank funding, higher liquidity and higher profitability supplied more credit. Credit supply seems to be also negatively correlated with a higher share of (gross) non-performing loans. This evidence is consistent with previous findings on the bank lending channel in Italy (di Patti and Sette (2016)) and in other countries (Khwaja and Mian (2008), Iyer et al. (2014), Jiménez et al. (2010)). The negative sign of bank capital is potentially counterintuitive, but this is not a novel finding in the literature (Berrospide

[7]See di Patti and Sette (2016) and Bofondi et al. (2017) for evidence of the impact on credit supply of the post-Lehman and the sovereign shocks, respectively, in Italy.

and Edge (2010) and Iyer et al. (2014)), suggesting that banks may choose higher capital rations because they are riskier, and therefore may lend less. Column 2 of Table 2.2 excludes mutual banks, which are subject to specific regulation on their operations and results are unchanged.[8]

2.3 The dynamics of aggregate productivity, its components, and aggregate credit supply

In this section we provide a brief sketch of the evolution of aggregate manufacturing productivity in Italy between 2003 and 2014, focusing on the driving forces that have shaped its dynamics, and proposing some suggestive evidence on its relationship with the fluctuations of credit supply. A comprehensive assessment of all these trends is offered in Figure 2.4, where the grayed out areas help identifying the periods of recession for the manufacturing sector.

Over the period of observation, the dynamics of value added in manufacturing has been particularly sluggish, experiencing a 9.3% drop between 2003 and 2014, as shown in panel (a). As a matter of fact, the sector experienced a recession in a third of the observed years, while not attaining a consistently fast-paced growth in the remaining ones. The massive drop in value added occurred in correspondence to the the global financial crisis (2007–09); after a modest rebound, it suffered a further —though more moderate— contraction during the sovereign debt crisis (2012–13).

The dynamics of manufacturing value added should be read in parallel to the chart displayed in panel (b), depicting the evolution of the aggregate credit supply shock. This has been obtained as an average of the credit supply shocks in equation 2.2, weighted by the share of loans granted in each sector and province. Credit supply has grown at rates above the mean until the global financial crisis. The growth of credit supply slightly increased in magnitude until 2006; after the outbreak of the crisis, the massive liquidity drought in interbank markets mirrored on the rapid shrinkage of credit supply growth. The pace of contraction slowed down in correspondence to the partial recovery of 2010, but another and more severe period of credit restriction was fostered by the sovereign debt crisis. A partial recovery emerged from 2013 on.[9]

[8]These banks have to allocate a minimum share of their loans to residents in the area in which they are headquartered, and are allowed to lend to shareholders.

[9]During the 2001–2003 recession, which didn't have a financial nature, the growth rate

How does the dynamics of aggregate labor productivity fit into these broad macroeconomic patterns? To provide a more insightful answer to this question, it is crucial to distinguish the role played by the reallocation of resources across firms from that played by the processes of firm entry and exit to/from the market.

2.3.1 The contribution of reallocation and firm demography

To quantify the relative contribution of different groups of firms to the dynamics of aggregate labor productivity, we exploit the decomposition proposed by Melitz and Polanec (2015). This is known as "dynamic Olley and Pakes decomposition", since it represents a dynamic extension of the widely-used decomposition used by Olley and Pakes (1996) to distinguish between the efficiency gains deriving from the reallocation of resources towards the most productive firms (measured by the so-called OP covariance term), and those arising from the productivity growth of individual firms (captured by average firm productivity).

Following Melitz and Polanec (2015), we define aggregate productivity as the average of firm-level log productivities, weighted by their share of employees. We then divide firms into three groups: entrants (E), exiting (X) and incumbent firms (S). Considering two consecutive time periods, it is possible to express the aggregate productivity of the first period (Φ_1) as the weighted average of the productivity of the firms that survive and the one of the firms that exit the market; analogously, the aggregate productivity of the second period (Φ_2) can be expressed as the weighted average of the productivity of the firms that survived and the one of the firms that have entered the market:

$$\Phi_1 = \Phi_{S1}\omega_{S1} + \Phi_{X1}\omega_{X1} \tag{2.3}$$

$$\Phi_2 = \Phi_{S2}\omega_{S2} + \Phi_{E2}\omega_{E2} \tag{2.4}$$

where Φ_{gp} is the aggregate productivity of group g in period p, and ω_{gp} is the share of employees in each group.

The difference between Φ_2 and Φ_1 returns the variation in aggregate productivity:

$$\Phi_2 - \Phi_1 = (\Phi_{S2} - \Phi_{S1}) + \omega_{E2}(\Phi_{E2} - \Phi_{S2}) + \omega_{X1}(\Phi_{S1} - \Phi_{X1}) \tag{2.5}$$

where the first term represents the productivity variation for the firms that are active on the market in both periods (the incumbents); the second is the

of credit supply remained constant and broadly in line with the levels displayed in 2004.

contribution of entrants, which is positive (negative) if their productivity is higher (lower) than the one of the incumbent firms; the third is the contribution of firms that exit the market, which is positive (negative) if their productivity is lower (higher) than the one of the incumbents.

Making use of the Olley and Pakes (1996) decomposition, the term $(\Phi_{S2} - \Phi_{S1})$ can be further decomposed in the variation of the incumbents' average productivity and the one of the covariance between incumbents' productivity and the share of employees, capturing the intensity of the reallocation process. To sum up, the variation of aggregate productivity can be expressed as the sum of the following four components:

$$\Phi_2 - \Phi_1 = \underbrace{\Delta\bar{\varphi}_S}_{\text{Avg. prod.}} + \underbrace{\Delta\text{Cov}_S}_{\text{Reallocation}} + \underbrace{\omega_{E2}(\Phi_{E2} - \Phi_{S2})}_{\text{Entry}} + \underbrace{\omega_{X1}(\Phi_{S1} - \Phi_{X1})}_{\text{Exit}} \tag{2.6}$$

How did these components evolve in our reference period? Going back to Figure 2.4, the dynamics of aggregate productivity —depicted in panel (a)— has been substantially similar to that of manufacturing value added, with wider fluctuations especially at the beginning of the sample. Reallocation —displayed in panel (c)— has always provided a positive contribution to aggregate labor productivity, partially offsetting the consistently negative contribution of average firm productivity (not reported in the figure, but available in Table 2.4). The contribution of reallocation moderately rose until 2007, and then momentarily slowed down, just before peaking in the wake of the two crisis episodes. It is interesting to note that the two jumps in the reallocation component seem to mirror the troughs experienced by credit supply, roughly with a lag of one year.

Panel (d) displays the contribution of entry and exit. The contribution of exiting (entering) firms is always positive (negative), since their aggregate productivity is always lower than the one of incumbents. The entry component fluctuates in a narrow band, just above the -2 percentage points, but displays no peculiar pattern. The exit component remains remarkably stable during the first part of our sample. After the global financial crisis, however, its contribution jumped up by 1 percentage point; it then appeared to converge back to its before-crisis values, but experienced another increase after the burst of the sovereign debt crisis. Like in the case of reallocation, the contribution of exiting firms also displays remarkable variations only in periods of substantial credit supply shrinkage.

Overall, this broad picture of productivity dynamics in Italian manufacturing provides some suggestive evidence of a causal link between the evolution of credit supply and certain components of aggregate labor productivity, most notably the reallocation and the exit terms. In the remainder of this paper, we exploit our firm-level data to provide some evidence in favor of this hypothesis, and to explore what are the mechanisms that give rise to the fluctuations we observe in the aggregate.

2.4 Firm-level evidence on the relevance of credit supply for productivity dynamics

In this section we exploit the ASIA dataset to investigate the effect of credit supply on firm behavior and performance, and on how this maps to the aggregate fluctuations documented in section 2.3. To guide our analyses, we will continuously make reference to the Melitz and Polanec (2015) decomposition discussed above, adopting regression models that speak as much as possible to the components of the aggregate productivity breakdown.

In its most general form, the specification adopted for most of the analyses presented in this section is the following:

$$y_{it} = \beta CSS_{sp,t-1} + \gamma X_{it} + \delta_{st} + \delta_{pt} + \varepsilon_{it} \tag{2.7}$$

where y_{it} is the firm-level dependent variable of interest; $CSS_{sp,t-1}$ is the credit supply shock, as defined in equation 2.2; X_{it} are firm-level controls; δ_{st} and δ_{pt} are a set of sector $\times$ time and province $\times$ time fixed effects, respectively; ε_{it} is an error term. Standard errors are clustered at the province and sector level to account for serial correlation.

The coefficient of interest is the one attached to the credit supply shock, which is included with a one-year lag, in order to allow for a delay in the transmission of the credit market shocks to the economic conditions of the firms.[10] Since the period spanned by our data includes a disruptive event such as the global financial crisis, followed in Italy by a further downturn as a consequence of the turmoil on the sovereign debt markets, we check if the effects of the credit supply shocks are heterogeneous across time, by splitting our dataset in two subsamples, containing information on the period before (until 2007) and during the crisis (from 2008 on).

Identification relies on the exogeneity of the credit supply shock with respect to the decisions and performance at the firm level. As alluded to in section 2.2, we claim this to be the case: on one side, the bank-specific shocks are by construction uncorrelated with any characteristics of the firms

[10] In the parsimonious specification presented in this section, we only explore the effect of a one-year-lag credit supply shock. We have explored more complex lag structures for the credit supply shock. The first lag always displays the larger and more significant effects; results are basically unchanged by including more lags.

and the markets in which the bank operate; on the other, the market shares used to aggregate the bank-specific shocks are fixed in time (according to the scheme described in equation 2.2), in order to avoid incorporating in our shock the banks' sectoral and geographic strategic positioning decisions, which could have been driven by the economic performance of firms within a given sector-province cell. In model 2.7, the sector $\times$ time and the province $\times$ time fixed effects control for sector-and province-specific time trends, and are intended to address additional concerns on omitted variables correlated to both the economic and the credit cycles.

Since the credit supply shock varies at the sector-province level, model 2.7 may be estimated on data collapsed at the sector-province-year level. However, we prefer to use firm-level data for two reasons. First, they allow us to investigate important phenomena, such as the reallocation of resources across firms, even within a certain sector-province cell. Second, we can exploit the available firm-level controls to augment the precision of our estimates. Firm-level controls, denoted as X_{it} in model 2.7, can include the age of the firm, the size class and the quintile of the productivity distribution it belongs to. Since size class captures most of the within-cell heterogeneity in firm outcomes, we typically use it as the sole firm-level control, unless stated otherwise. However, results are robust to the inclusion of other controls.

In the remainder of this section, we will separately analyze the effect of credit supply on three different groups of firms: exiting, entering and incumbents.

2.4.1 Incumbents

In the case of incumbents, the firm-level counterparts of the decomposition in equation 2.6 is straightforward. The average productivity (or within-firm) term and the reallocation one can be rewritten, respectively, as:

$$\Delta\bar{\varphi}_{S,t} = \frac{1}{N_t}\sum_{i\in S}(\varphi_{i,t} - \varphi_{i,t-1}) \tag{2.8}$$

$$\Delta \mathrm{Cov}_{S,t} = \sum_{i\in S}\left[\varphi_{i,t}\left(s_{i,t} - \frac{1}{N_t}\right) - \varphi_{i,t-1}\left(s_{i,t-1} - \frac{1}{N_t}\right)\right] \tag{2.9}$$

where N_t is the number of incumbents and $s_{i,t}$ is the share of employees, computed over the whole manufacturing sectors, working for firm i at time t.

We can therefore capture the effect of a credit supply shock on these two components, by using the average productivity (2.8) and the reallocation (covariance) component (2.9) as dependent variables in our regression framework. In the case of equation 2.8, this boils down to using the log growth rate of labor productivity at the firm level. As for equation 2.9, note that the components of the covariance are yearly changes in the firm level productivity weighted by its share of employees.[11]

Table 2.5, panel (a), displays estimates of model 2.7 on labor productivity growth at the firm level. The credit supply shock has a positive and significant coefficient; as columns (2) and (3) show, the result is entirely driven by the years during the crisis, mostly characterized by negative supply shocks. The sign of this relationship may reflect both managerial choices at the firm level (for example, through the dynamics of investments) and short-run adjustments in sales that are not accompanied by contemporaneous adjustments in terms of employees. In terms of magnitude, the effect is sizable but not huge: a one-standard-deviation increase of the credit supply shock raises the growth rate of firm-level labor productivity by 0.8 percentage points, equal to one twentieth of a standard deviation. These elasticities are in the same order of magnitude as the ones estimated by other studies focusing on the effects of credit supply shocks on Italian firms: Manaresi and Pierri (2016), using firm-level TFP and a sample of significantly larger firms, estimate elasticities ranging between 0.11 and 0.16, depending on the specification.

The results in panel (b) of Table 2.5 use instead the firm-level contribution to reallocation as a dependent variable. In this case, the credit supply shock has a negative coefficient, again driven by the years during the crisis. This indicates that the positive contribution of reallocation to productivity growth is stronger in those sector-provinces that experienced a relatively stronger fall in credit. The point estimates are very small in size, reflecting the scale of the dependent variable. However, the economic relevance of this effect has to be evaluated in the aggregate, since the individual contributions enter additively and are not averaged out in the aggregation process, as shown in equation 2.9. For an average number of incumbents of 410 thousand units, a uniform one-standard-deviation increase of the credit supply term would make the aggregate reallocation component drop by roughly 2 percentage points.

[11] The summands somehow resembles the growth rate of productivity at the firm level, though now the two log terms are weighted by their share of employees relative to the mean (which equals $1/N$); for a firm with a share of employees above the average, this term is positive either if the firm increases its productivity or if it further gains employment shares. Rearranging terms, the summand can alternatively be interpreted as the difference between the firm-level contribution to aggregate productivity and the firm-level contribution to average productivity; if the former is greater than the latter, the firm positively contributes to reallocation.

To provide further support to the evidence on the sizable effect of credit supply shocks on reallocation and shed some light on the underlying mechanisms, we test the effect of the credit supply shock on the growth rate of employees at the firm level. Table 2.6 shows a positive and significant coefficient, once more driven by the years during the crisis. This tells us that —while in good times credit shocks do not on average induce firm to grow in size— during credit restrictions firm release some of their employees. This effect is not uniform across firms: in the set of results in the bottom panels of the table, we allow the effect of credit supply to be heterogeneous across different quintiles of the productivity distribution and size classes. Results show that the effect of the credit supply shock is concentrated in the lowest quintiles of the productivity distribution, while it doesn't have a significant impact on more productive firms. This effect is entirely driven by the observations belonging to the period during the crisis: when a negative credit shock hits the economy, less productive firms will on average reduce their employment, while the better performing ones will not modify their scale. We observe a similar pattern when we look for differential effects across size classes: in times of credit restriction, small- and medium-sized firms (up to 50 employees) reduce their workforce, while no significant effects emerge for bigger ones.[12] Interestingly, in good times most of the action takes place in the upper tail of the firm size distribution, with bigger firms growing in size in response to a positive supply shock.

This evidence, which is consistent with a more selective economic environment arising as a consequence of a credit crunch, suggests that the credit restrictions experienced by Italian firms during the recessions of 2009 and 2012–2014 have had a "cleansing effect" on manufacturing. This has implied a redistribution of employment shares (though not of employees) in favour of bigger and more productive firms, and is therefore in line with a more prominent role of reallocation on the dynamics of aggregate productivity during a credit restriction.

[12]The effect we find on reallocation might be exacerbated by a specific feature of Italian labor market that we are not able to control for. The Redundancy Fund (*Cassa Integrazione Guadagni*) is an institute that is intended to provide support to firms in a temporary situation of difficulty, financing the wage bill of the workforce unused in the production process. During the crisis, this instrument has been extensively used by the entitled firms (those above 15 employees), which in this way were able to delay the time of the workers' layoff. This can partly explain why we find an effect mostly on the firms belonging to lower size classes; if that was the case, the reallocation of employment shares from small to big firms would be "artificially" inflated (since workers benefiting from the Redundancy Fund do not actively take part in production). Unfortunately, our data do not allow us to single out the workers that benefit from Redundancy Fund.

2.4.2 Exiting firms

Going back to the decomposition in equation 2.6, let's have a closer look to the term that pins down the contribution of exiting firms to the dynamics of aggregate productivity. It is equal to the product between the share of employees belonging to exiting firms and the relative productivity of incumbents over exiting firms (expressed in terms of log difference). Since the productivity of exiting firms is always likely to be lower than that of the incumbents, we expect this term to be positive. Therefore, the higher the share of workers employed by exiting firms and the lower their productivity relative to the incumbents, the higher will be the contribution of exit to the growth of aggregate productivity. Additional insights can be gained by further decomposing the exit term as follows:

$$\text{Exit}_t = \omega_{Xt}(\Phi_{St} - \Phi_{Xt}) = \underbrace{\frac{N_{Xt}}{N_t}}_{\text{Exit rate}} \times \underbrace{\frac{L_{Xt}/N_{Xt}}{L_t/N_t}}_{\text{Relative size}} \times \underbrace{(\Phi_{St} - \Phi_{Xt})}_{\text{Relative productivity}} \tag{2.10}$$

where L and N denote the number of employees and the number of firms, respectively. This formulation highlights that —apart from the already discussed relationship with relative productivity— the contribution of exit is positively related to both the exit rate and the relative size of exiting firms. Taking this decomposition to our data, we are able to disentangle the driving forces behind the observed increase in the exit component during the two episodes of crisis. Figure 2.5 displays the evolution of the three terms, together with the overall evolution of the exit component of aggregate productivity, that we report again in panel (a). The substantial increase observed after the burst of the global financial crisis has been fostered by both the rise in the exit rate —which peaked in correspondence with the two crisis events— and the steady deterioration of exiting firms' productivity with respect to the incumbents' one.[13] The overall effect has been only partially limited by the reduction in exiting firms' relative size, which was already underway before the crisis.

We now try to exploit the firm-level information in our data to gain some insight on the effect of credit supply on these aggregates. To do that, we apply model 2.7 to firm-level counterparts of the three terms singled out in equation 2.10.

[13] Panel (d) is the exact counterpart of the last term in equation 2.10, displaying the log difference between the aggregate productivity of incumbents and of exiting firms. Hence, an increase in this series implies that exiting firms become progressively less productive than incumbents.

For the exit rate it is quite straightforward to use a linear probability model, taking as the dependent variable a dummy equal to one if firm exits in year t; indeed, the conditional expectation of this model returns the aggregate exit rate displayed in panel (b) of Figure 2.5. Regression results are shown in Table 2.7, panel(a). As expected, the probability of exit is negatively related to past credit supply shocks. The effect is driven by the period after the outbreak of the crisis, which is mostly characterized by negative shocks: after 2008, a decrease of the credit supply shock by one standard deviation makes the exit rate increase by 0.5%, which can be roughly quantified as an expulsion of 2,500 firms from the market. As column (2) shows, credit supply does not significantly affect the probability of exit in normal times.

To document the response of relative size to a credit shock, we restrict our sample to exiting firms only, and we take as a dependent variable the ratio of the employees in firm i over the average number of employees per firm in each sector-province cell. That is, denoting with the index $i(X)t$ all the firms i that exit the market at time t, we define our dependent variable as $L_{i(X)t}/(L_{spt}/N_{spt})$. The results displayed in panel(b) of Table 2.7 document a positive relationship between credit supply and relative size of exiting firms; the magnitude of the effects in the two subperiods is similar, although the coefficient obtained on the before-crisis sample turns out to be marginally non-significant. Under a one-standard-deviation contraction of credit supply, the relative size of an exiting firm shrinks by is 3.5 percentage points, roughly the 8% of the average relative size of exiting firms in our sample period.

We keep focusing on exiting firms only, to investigate the response of relative productivity. In this case, the dependent variable is the difference between the log aggregate productivity of incumbents within each sector-province cell and the log productivity of each exiting firm; for each firm i that exit the market at time t, we therefore define it as $\Phi_{Sspt} - \varphi_{i(X)t}$, where Φ_{Sspt} is the aggregate productivity of the stayers at the sector-province level. As in the case of relative size, it is the firm level counterpart of the term listed in equation 2.10, computed with respect to the sector-province cell it belongs to. Results are shown in Table 2.7, panel(c). The relative productivity term is negatively related to the credit supply shock, meaning that exiting firms become relatively less productive when hit by a negative shock. Columns (2) and (3) show that the sign of the effect does not differ before and after the crisis, although the magnitude is higher in the latter subsample.

These results indicate that a negative credit shock induces a drop in size and productivity of exiting firms, relative to stayers. At least two forces are driving this result: (i) everything else equal, it can be expected that during a credit crunch the level of employees/productivity deteriorates faster for exiting firms than for continuing ones; (ii) if a credit crunch fosters a greater selection on the market, the firms that manage to survive will be increasingly bigger/more productive in relative terms. Since we are only observing firms in the year of their exit, these patterns are not informative of the intrinsic quality of the exiting firms; as a matter of fact, we would expect a period of credit crunch to force out of the market relatively better firms, both in terms of size and in terms of productivity (Caballero and Hammour, 1994; Foster et al., 2014).

To provide further evidence on the selection induced by a negative credit shock on exiting firms, we regress some variables that capture firm performance in the three years before the exit (log productivity, log employees, and the growth rate of both of them) on a dummy equal to one if the firms exits during a credit crunch.[14] All controls —apart from the credit supply shock— are the same as in model 2.7. Importantly enough, we include a second-order polynomial in the firm age to rule out the possibility that our dummy selects the longer-surviving firms, since the negative shocks are concentrated towards the end of our sample. Results are shown in Table 2.8, and confirm that —when compared to the firms that exit in credit expansion times— firms exiting during a credit crunch are on average bigger, more productive, and display a faster growth of employees and productivity.

2.4.3 Entering firms

Like the exit term, the entry one can also be decomposed in the following way:

$$\text{Entry}_t = \omega_{Et}(\Phi_{Et} - \Phi_{St}) = \underbrace{\frac{N_{Et}}{N_t}}_{\text{Entry rate}} \times \underbrace{\frac{L_{Et}/N_{Et}}{L_t/N_t}}_{\text{Relative size}} \times \underbrace{(\Phi_{Et} - \Phi_{St})}_{\text{Relative productivity}} \tag{2.11}$$

The entry contribution is expected to be negative, since the average entrant is relatively less productive than incumbents.

Figure 2.6 shows the dynamics of the three terms in equation 2.11 and of the overall entry component of aggregate productivity. The entry contribution displays wide fluctuations, mostly driven by the swings of the relative productivity of entrants with respect to incumbents, which seems to peak in correspondence to the most severe years of crisis.[15] This is consistent with

[14] When performing this exercise, we somehow improperly refer to "firms exiting during a credit crunch". As a matter of fact, it would be more correct to talk about firms than enter in times of credit supply growth below the sample mean.

[15] Extending the time span of our data to the previous crisis episode (2002–03, not used in this paper because of data incompleteness) reveals another peak in relative productivity in correspondence with the years of downturn.

the evidence provided by Lee and Mukoyama (2015), who show that the firms entering during a recession are on average more productive. The steadily-declining entry rate has dampened the negative contribution of entry; the relative size of entrants has instead operated in the opposite direction, with a sudden increase —although limited in size— in correspondence to the years of the crisis.

The empirical analysis is almost entirely symmetrical to the one performed for exiting firms.

The most notable difference rests in the analysis of the entry rate. Since we do not observe the pool of potential entrants before their entry into the market, estimating a linear probability model on entrants would make little sense. As a consequence, in this case we compute the entry rate at the sector-province level (N_{Xt}/N_t) and run model 2.7 on collapsed data, dropping firm-level controls. The results presented in panel (a) of Table 2.9 show that the credit supply shock does not significantly affect the entry rate.

Going back to data at the firm level, and limiting the sample to entrant firms only, we then analyze the response of relative firm size. This is now computed as $L_{i(E)t}/(L_{spt}/N_{spt})$, where the firms indexed by $i(E)t$ are those that enter the market at time t. The results displayed in panel (b) of Table 2.9 show that a credit supply restriction implies a drop in the relative size of entrants. The estimated intensity of the effect does not differ across subsamples, and the size of the coefficients is roughly comparable to the one estimated for exiting firms.

The relative productivity of entrants —this time defined as the difference between the log productivity of each entrant and the log aggregate productivity of incumbents within each sector-province cell ($\varphi_{i(E)t} - \Phi_{Sspt}$)— is positively related to the credit supply shock, as shown in Table 2.9, panel (c). A one standard deviation contraction of credit supply would imply a drop of about 10 percentage points in the relative productivity of entrants. The elasticity is lower in the years before the crisis.

To provide some intuition on the nature of the firms entering the market during a credit restriction, we compare the performances of the firms for which we can observe the entry year. Restricting the sample to these firms and focussing on the three years after entry, we regress their performance in terms of employees and productivity growth on a dummy equal to one if the firms has entered during a credit crunch. All controls are the same as those used to obtain the results in Table 2.8. According to the estimates shown in Table 2.10, firms entering during a credit crunch do not display significant differences in terms of size; on the other hand, they are on average more productive and they display a faster growth of employment and productivity. Moreover, their exit rate at a three-year distance from the entry (column (5)) is significantly lower than the one displayed by firms entered during credit expansion times.

Based on the set of results presented so far, in the next section we will push forward an exercise of aggregation intended to quantify the overall effect of the credit crunch suffered by Italian firms after the crisis, isolating the individual channels that —with different signs and intensities— contributed to shape the dynamics of Italian manufacturing labor productivity.

2.5 Aggregate implications

In this section we perform a back-of-the-envelope calculation of the aggregate effects that the credit restriction registered after the burst of the crisis has had on the dynamics of labor productivity in the Italian manufacturing sector.

We first define a counterfactual situation, in which the credit supply continued to grow at the same rate observed before the outbreak of the global financial crisis. To simulate this hypothetical scenario, we replace all the credit supply shocks after 2008 with the average credit supply growth observed before the crisis at the sector-province level. In other words, the evolution of the credit supply shock under this scenario would be:

$$\widetilde{CSS}_{spt} = \begin{cases} CSS_{spt}, & \text{if } t \leq 2007 \\ \frac{1}{5}\sum_{t\in[2003,2007]} CSS_{spt}, & \text{if } t > 2007 \end{cases} \tag{2.12}$$

As argued in section 2.2, the credit supply shock cannot be given an absolute quantitative interpretation, since it is identified net of a constant scaling factor. As a consequence, it would have little economic sense to stick to a particular value of the shock when building our hypothetical scenario, since we wouldn't be able to determine whether the imposed value corresponds to an expansion or a contraction of the credit supply. Our strategy of relying on the average pre-crisis shocks allows us to circumvent this problem, since (i) even though we cannot give it an absolute interpretation, the estimated shock still captures the relative differences between credit supply dynamics across sector-provinces and over time: this allows us to state that, on average, our hypothetical situation implies a 14 percentage points higher dynamics of the credit supply with respect to the one that was actually observed; (ii) the average credit supply shock is higher before than after the crisis in virtually all of the sector-province cells. Hence, although we are not able to assert that our counterfactual scenario implies an expansion of the credit supply in all the sector province cells, we can state that it represents a substantial improvement of the credit supply dynamics with respect to the realized ones.

With the counterfactual shock as expressed in equation 2.12 in hand, we go back to the results of the regressions presented in section 2.4, and for each of them we compute the predicted values at the firm level under the counterfactual scenario.[16] Aggregating these predicted values with the appropriate weights allows us to recover the dynamics of each term of the Melitz and Polanec (2015) decomposition, under the assumption of the credit supply expanding at the average pre-crisis rate.

This is of course a rough estimate, since aggregation weights are also a function of the credit supply shock, whose variation is not accounted for under this setup. However, the exercise is informative on the direction that the underlying determinants of aggregate productivity would have taken in the absence of a credit crunch, and provides an approximate estimate of their relative importance in the wake of the crisis.

Based on the results presented in Table 2.5, we first perform this exercise on the incumbents. Figure 2.7 shows the actual evolution of the terms in equation 2.10 (solid line), as opposed to those that would occur in the counterfactual scenario (dashed line). The credit restriction after 2008 has negatively affected the dynamics of within-firm productivity: with respect to the counterfactual scenario, the contribution of average productivity has been lower by 1.2 percentage points per year in the period 2008–2014. On the contrary, the 2008–2014 credit crunch has fostered a higher contribution of reallocation to aggregate productivity, quantifiable in roughly 3.5 percentage points per year on average.

Figure 2.8 shows the results of the same exercise applied to the exit component of aggregate productivity. Overall, the credit crunch has sustained the exit channel of aggregate productivity dynamics: its cumulative contribution can be quantified in roughly 3 percentage points over the 2008–2014 period. This result is driven both by the exit rate and by the relative productivity of incumbents with respect to exiting firms, that would have been substantially lower if the credit supply had kept growing at the pre-crisis rate. By contrast, the relative size of exiting firms would have been slightly higher in the counterfactual scenario.

As for entrants, the results are displayed in Figure 2.9. In this case, the credit crunch has weighed negatively on the entry channel of aggregate productivity dynamics: over the whole period 2008–2014, its cumulative contribution is negative by about 1.5 percentage points. The result is driven by the dynamics of both the relative productivity and the relative size of entrants, which would have been higher under the hypothesis of no credit crunch. En-

[16] To perform this exercise, we adopt the more conservative parameters estimated on the whole sample (that is, the ones contained in column (1) of each table).

try rate shows instead little variation in the counterfactual scenario, as one could expect given the estimates shown in Table 2.9.

In Figure 2.10 we provide an overall evaluation of the impact of the credit crunch started in 2008, separately looking at incumbent firms (the sum of average productivity and reallocation) and firm demography (the sum of entry and exit). The tighter credit conditions have as a whole sustained the dynamics of aggregate manufacturing productivity: as panel (a) shows, if the credit supply had expanded at the same pace it did before the crisis, the growth rate of the incumbents' aggregate productivity would have been lower by 2.4 percentage points per year over the period 2008–2014. As panel (b) of Figure 2.10 shows, credit restriction has sustained the dynamics of aggregate productivity through firm demography as well, meaning that the exit channel has dominated the entry one; the effect is substantially smaller, though, totalling additional 0.2 percentage points per year. This results are driven by the reallocation and exit components, whose positive contribution has increased, benefiting from the more selective economic environment arising as a consequence of the credit tightening. Overall, the rise in these two components has more than compensated the negative effect that the credit crunch has had on within-firm productivity and entry.

These findings suggest that credit shock contribute to the cleansing effect of recessions. During both the financial and the sovereign crises in Italy, they forced low productivity incumbent firms to shrink, in this way contributing positively to the growth of aggregate productivity. However, this has come at the cost of a substantial fraction of the workforce being laid off, as a consequence of the credit restriction. If we replicate a similar aggregation exercise using the results on the growth rate of employees displayed in Table 2.6, we can quantify that roughly 100,000 jobs were destroyed among the incumbent firms only, as a consequence of the 2008–14 credit crunch; these numbers correspond to the 12% of the observed reduction in manufacturing employment over the same period. If the workers laid off by low productivity firms are not reabsorbed, there might be negative long-term effects on the productivity of the workforce, as well as important welfare costs. We plan to quantify to what extent the reallocation effects we identified translate into higher unemployment in future work.

2.6 Conclusions

In this paper we study if and to what extent credit supply shocks can account for fluctuations in aggregate labor productivity. To isolate the different channels through which credit supply affects productivity, we base our empirical approach on the decomposition proposed by Melitz and Polanec (2015), which breaks down the dynamics of aggregate productivity into four components: the variation of average firm productivity, the reallocation of resources towards more productive firms, the contribution of exit and the contribution of entry. Closely following this interpretation framework, we exploit a unique dataset on the universe of Italian manufacturing firms to study the impact of a credit supply shock at the industry-province level on each of these components. We isolate credit supply shocks applying the procedure proposed in Greenstone et al. (2014) on detailed microdata from the Italian Credit Register.

The results of the decomposition show that the sluggish aggregate manufacturing productivity in Italy in the period 2003–2014 is primarily driven by the negative contribution of the average (within-firm) productivity. Reallocation of resources to more productive firms has instead sustained the dynamics of aggregate productivity in all years, though its relevance spiked during the global financial and the sovereign debt crises, which were characterized by a massive restriction of credit supply. The exit component of the productivity decomposition, which always contributes positively (since on average exiters are less productive than incumbents), increased in magnitude after 2009, too.

This evidence, suggesting that credit supply shocks may reverberate on aggregate productivity through various channels, has been more rigorously explored in a regression framework that exploits firm-level data to closely track the components of the Melitz and Polanec (2015) decomposition. Our findings show that a restriction in credit supply affects aggregate productivity growth mainly through the reallocation component: less productive firms shrink in size as a consequence of a negative credit supply shock, thus losing employment shares in favour of more productive ones. Productivity growth benefits by a credit restriction also through the exit component, whose relevance rises as a consequence of the increase in the exit rate and the reduction in relative productivity of exiters with respect to incumbents. On the other hand, a negative credit supply shock hinders aggregate productivity growth through other channels, such as the within-firm productivity (because of the lower productivity growth of the incumbents) and the entry component (because firms entering the market are smaller and less productive than incumbents).

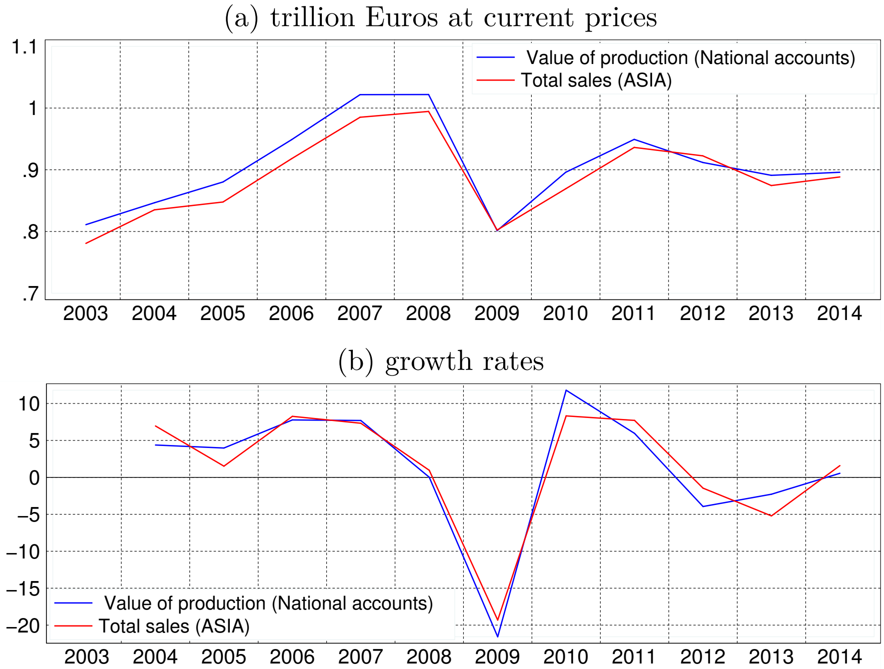
Since our sample encompasses both a period of tranquil credit market conditions and a period of severe credit restriction, it is natural to ask what has been the relative importance of these channels in the wake of the double recession that hit Italy in the period 2008–2014. To this aim, we set up a counterfactual scenario under which —in each sector and province— credit supply continued to grow at the same average rate recorded in the pre-crisis period. We then apply the estimates obtained in our firm-level regressions to this hypothetical situation, to provide a rough quantification of the role played by credit restriction in shaping the growth pattern of aggregate productivity. We find that, if after 2008 credit supply had continued to grow at its pre-crisis rates, the growth rate of aggregate productivity would have been lower by roughly 2.5 percentage points per year. This discrepancy would have been primarily driven by reallocation and —to a lesser extent— by the exit channel, whose negative contribution would have more than offset the positive effect on within-firm productivity and entry.

When interpreting the results of this paper, it is important to keep in mind that aggregate productivity is not an all-encompassing welfare measure. As an example, let's compare the situation in which only one small and very productive firm remained on the market with the one in which many firms —spanning the whole range of productivities— employed a substantial fraction of the workforce: aggregate productivity would be higher in the former case, though the latter would probably be more socially desirable. In this paper we document the impact of the credit crunch on the dynamics of manufacturing aggregate productivity in Italy; nonetheless, this impact comes at the cost of an increase in firms' mortality and a drop in employment (concentrated in the firms at the lower end of the productivity distribution). According to our estimates, the credit restriction observed in the 2008–14 period is responsible for the destruction of about 100,000 jobs among the incumbent firms only, which are the 12% of the observed drop in manufacturing employment over the same period. Providing a broader assessment of the effects of the credit crunch would at least imply looking at the reallocation of workers and economic activity from the manufacturing sector to other sectors of the economy, which we plan to do in future work.

Finally, our findings indicate that most of the gains come from the reallocation of employment shares to more efficient firms. However, the relevance of this channel could be especially large in a country like Italy, which is characterized by a high level of misallocation (Calligaris et al., 2016; Gamberoni et al., 2016) and therefore present a greater scope for reallocation. The positive effects of the credit restriction on reallocation may be smaller in other countries.

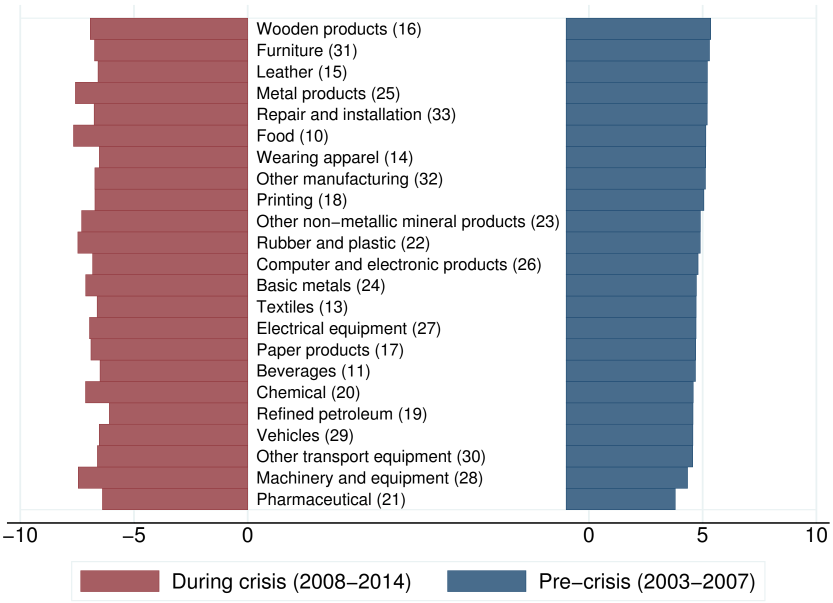
Figures and tables

Figure 2.1: Comparison between National Accounts and ASIA dataset



Source: National accounts and ASIA database.

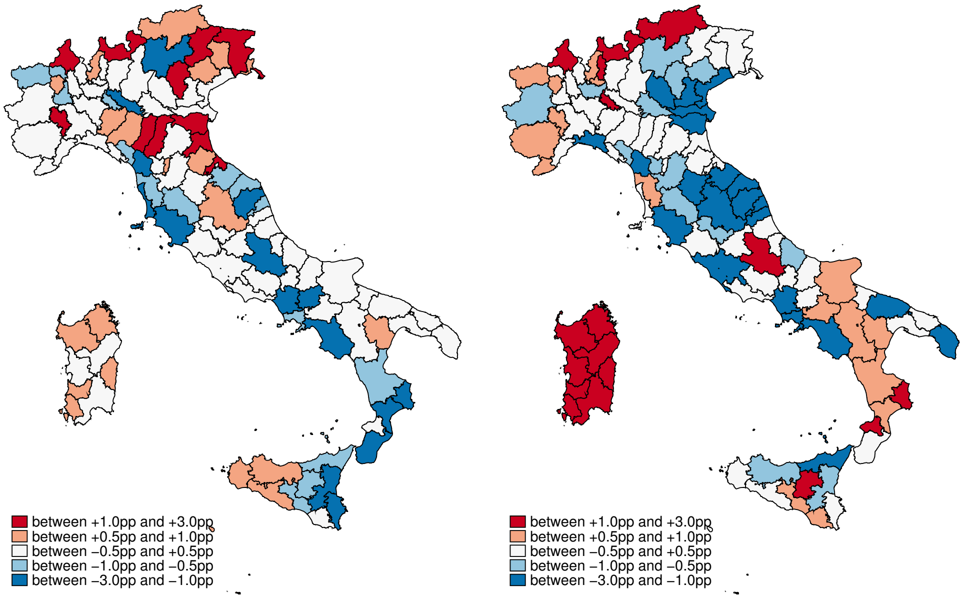
Figure 2.2: The credit supply shock, by sector



Source: Own elaborations from Italian Credit Register data.

Figure 2.3: The geography of credit supply shocks

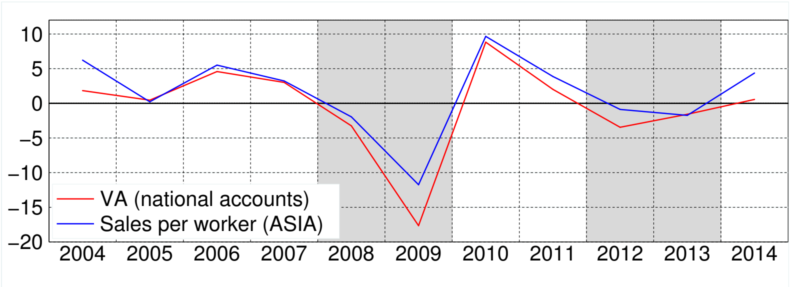
(a) Before the crisis (2003–2007) (b) During the crisis (2008–2014)



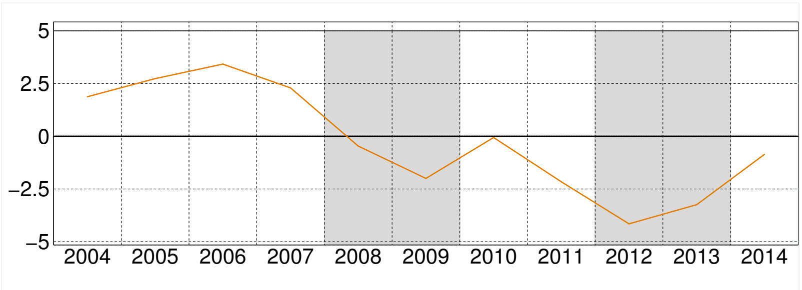
Source: Own elaborations from Italian Credit Register data.

Figure 2.4: Italian manufacturing, growth rates, 2004–2014

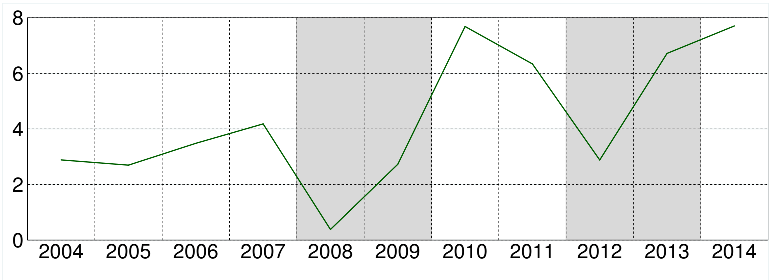
(a) aggregate VA and productivity (sales per worker)



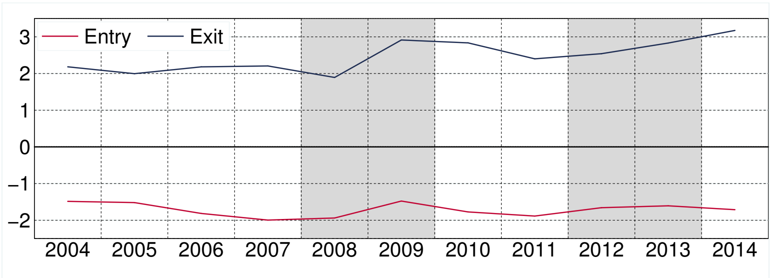
(b) aggregate credit supply



(c) contribution of reallocation to aggregate productivity growth

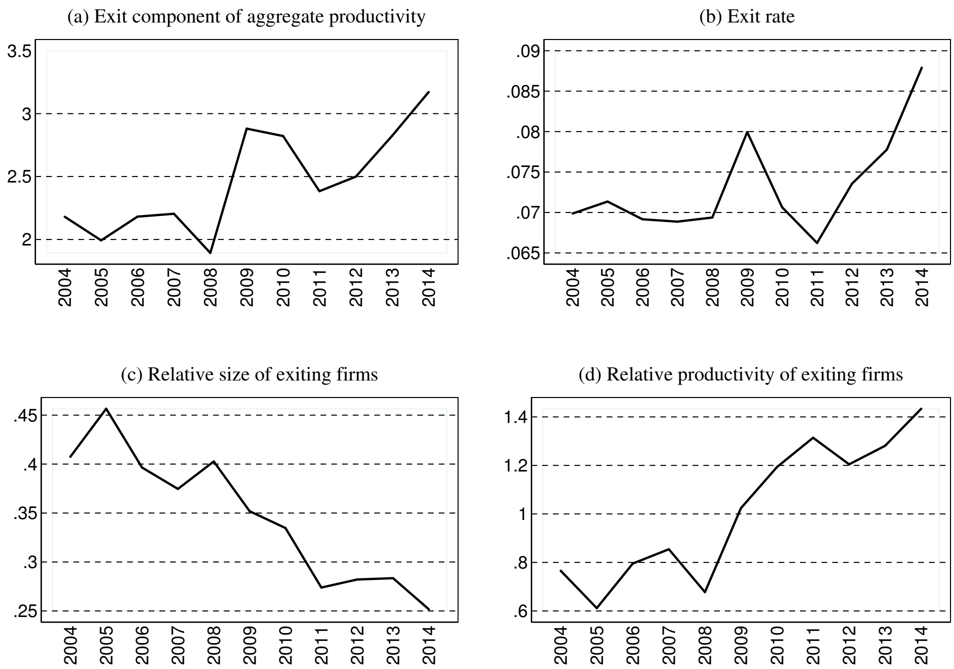


(d) contribution of entry & exit to aggregate productivity growth



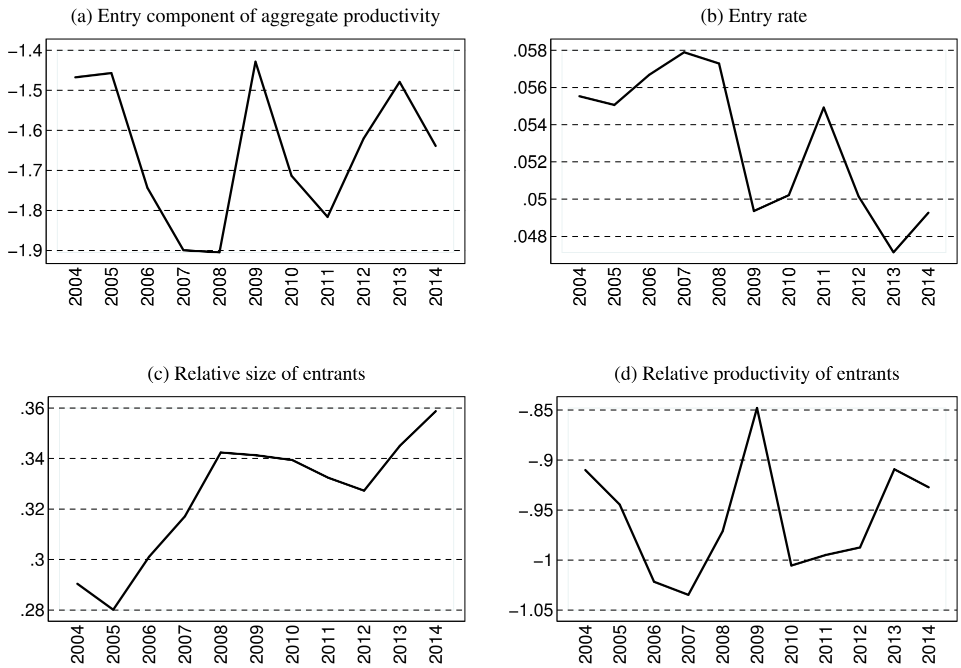
Notes: National accounts and own elaborations from Italian Credit Register and ASIA databases. Grayed out areas correspond to years of recession for the manufacturing sector.

Figure 2.5: Exit component of aggregate productivity and its determinants



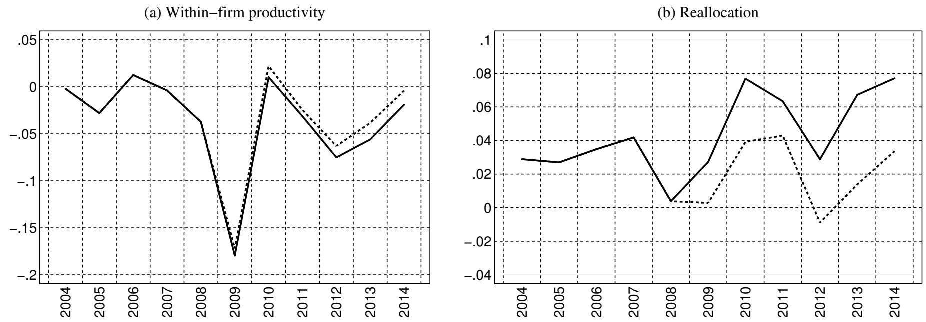
Notes: The exit component of aggregate productivity —depicted in panel (a)— is equal to the product of the variables in the remaining panels. Relative size is expressed as the ratio between the average number of employees in exiting vs. incumbent firms. Relative productivity is the log difference between the aggregate productivity of incumbents and the one of exiting firms.

Figure 2.6: Entry component of aggregate productivity and its determinants



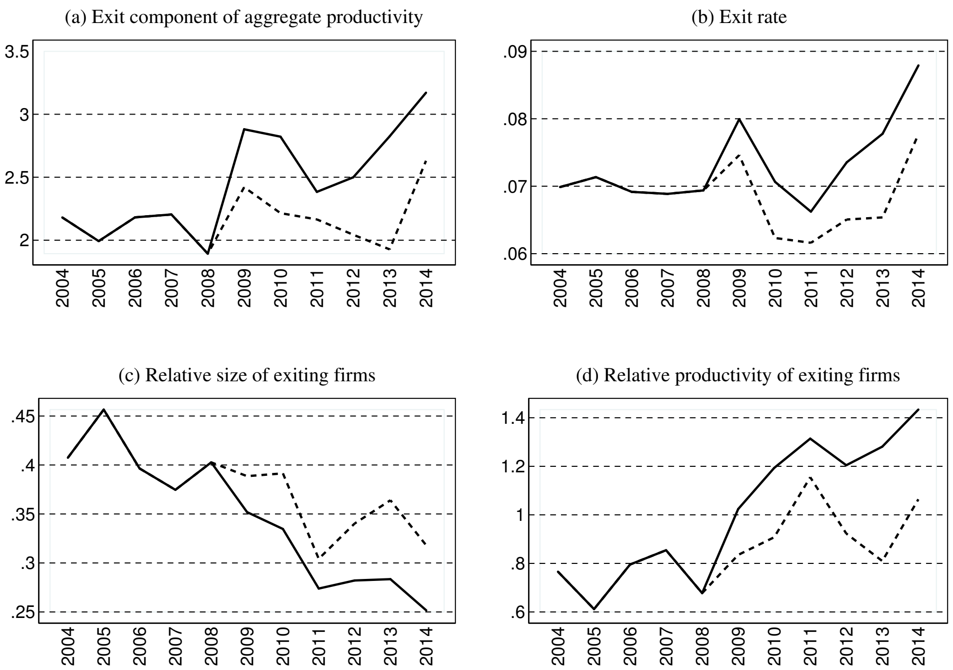
Notes: The entry component of aggregate productivity —depicted in panel (a)— is equal to the product of the variables in the remaining panels. Relative size is expressed as the ratio between the average number of employees in entrant vs. incumbent firms. Relative productivity is the log difference between the aggregate productivity of entrants and the one of incumbents.

Figure 2.7: Quantifying the effect on entry of the 2008–2014 credit crunch - Incumbents



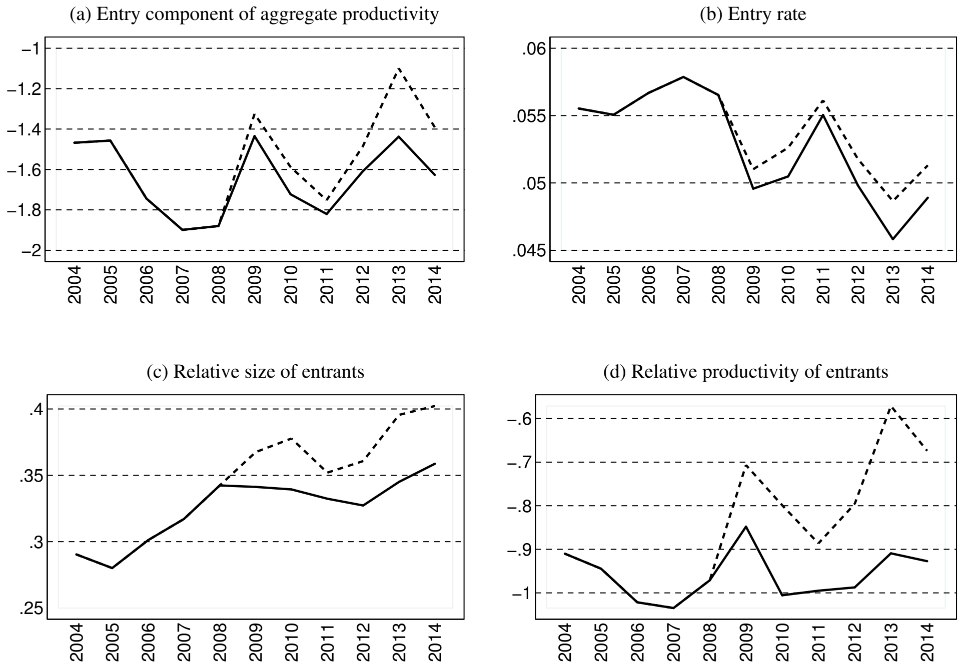
Notes: The dashed lines represent the predicted dynamics of each component, when we set the credit supply shocks equal to their pre-crisis averages from 2008 on.

Figure 2.8: Quantifying the effect on exit of the 2008–2014 credit crunch - Exiting firms



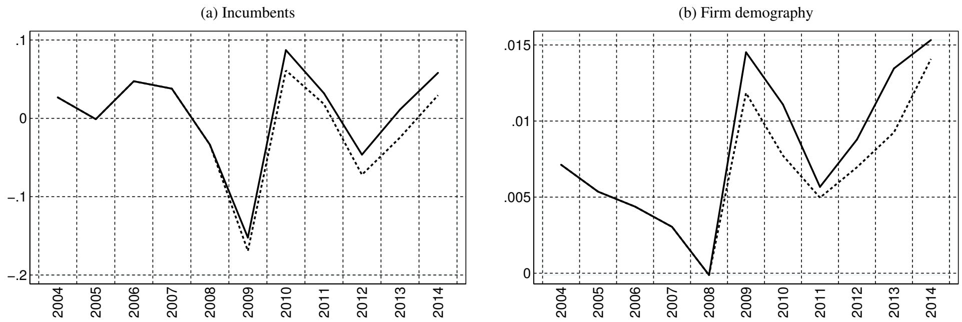
Notes: The dashed lines in panels (b)–(d) represent the predicted dynamics of each component, when we set the credit supply shocks equal to their pre-crisis averages from 2008 on. The dashed line in panel (a) is obtained as a product of the other ones. Relative size is expressed as the ratio between the average number of employees in exiting vs. incumbent firms. Relative productivity is the log difference between the aggregate productivity of incumbents and the one of exiting firms.

Figure 2.9: Quantifying the effect on entry of the 2008–2014 credit crunch - Entrants



Notes: The dashed lines in panels (b)–(d) represent the predicted dynamics of each component, when we set the credit supply shocks equal to their pre-crisis averages from 2008 on. The dashed line in panel (a) is obtained as a product of the other ones. Relative size is expressed as the ratio between the average number of employees in entrant vs. incumbent firms. Relative productivity is the log difference between the aggregate productivity of entrants and the one of incumbents.

Figure 2.10: Quantifying the effect on aggregate productivity of the 2008–2014 credit crunch



Notes: The dashed lines represent the predicted dynamics of each component, when we set the credit supply shocks equal to their pre-crisis averages from 2008 on.

Table 2.1: Descriptive statistics for manufacturing, years 2000–2014

	# firms	# employees	avg. size	sales	sales p.w.
			Levels		
2003	497,751	4,551,915	9.14	898,515	197,393
2004	487,815	4,466,044	9.16	936,484	209,690
2005	482,369	4,411,785	9.15	927,168	210,157
2006	477,894	4,395,526	9.20	974,697	221,748
2007	473,469	4,432,864	9.36	1,014,716	228,908
2008	459,217	4,388,661	9.56	984,992	224,440
2009	438,678	4,153,744	9.47	822,789	198,084
2010	426,504	4,001,394	9.38	869,212	217,227
2011	425,312	3,982,285	9.36	898,559	225,639
2012	417,228	3,897,932	9.34	871,785	223,653
2013	407,307	3,782,829	9.29	831,344	219,768
2014	396,401	3,704,193	9.34	849,658	229,377
			Growth rates		
2004	-2.00	-1.89	0.11	4.23	6.23
2005	-1.12	-1.21	-0.10	-0.99	0.22
2006	-0.93	-0.37	0.56	5.13	5.52
2007	-0.93	0.85	1.79	4.11	3.23
2008	-3.01	-1.00	2.08	-2.93	-1.95
2009	-4.47	-5.35	-0.92	-16.47	-11.74
2010	-2.78	-3.67	-0.92	5.64	9.66
2011	-0.28	-0.48	-0.20	3.38	3.87
2012	-1.90	-2.12	-0.22	-2.98	-0.88
2013	-2.38	-2.95	-0.59	-4.64	-1.74
2014	-2.68	-2.08	0.62	2.20	4.37

Notes: Own elaborations from ASIA dataset. Sales data are expressed in million Euros. Both sales and sales per worker have been deflated to 2010 values. Average size is expressed in terms of employees per firm.

Table 2.2: Credit supply shocks and bank balance-sheet conditions

	(1) All banks	(2) Non-mutual banks
Total capital	-0.00049***	-0.00058***
	(0.00018)	(0.00019)
Interbank exposure	-0.00112***	-0.00139***
	(0.00036)	(0.00044)
Liquidity ratio	0.00149***	0.00067
	(0.00021)	(0.00043)
ROA	0.01369***	0.01076**
	(0.00419)	(0.00410)
NPL/total asset	-0.00418***	-0.00948***
	(0.00120)	(0.00239)
Log of assets	-0.00398**	-0.00702***
	(0.00154)	(0.00213)
Dummy BCC banks	-0.01856***	
	(0.00457)	
Obs.	6,699	1,789
R^2	0.3196	0.2361

Source: Own elaborations from Italian Credit Register data. *Notes:* All regressions include year fixed effects. Clustered standard errors in parentheses: * p<0.10, ** p<0.05, *** p<0.01. Results are robust to the inclusion of different bank capital measures, such as the capital to assets ratio, or the Tier1 ratio.

Table 2.3: Summary statistics for the credit supply shock

	Mean	Std. dev.	25^{th} pct.	Median	75^{th} pct.
2003	3.854	3.446	1.998	3.470	5.302
2004	4.079	3.122	2.217	3.936	5.748
2005	7.210	2.672	5.898	7.432	8.712
2006	9.593	3.123	8.154	9.666	11.206
2007	6.492	3.579	4.419	6.443	8.461
2008	-3.202	5.391	-6.036	-3.537	-0.909
2009	-7.714	4.164	-10.288	-8.041	-5.419
2010	-1.498	2.808	-2.486	-1.423	-0.380
2011	-7.486	2.996	-9.104	-7.541	-5.868
2012	-14.525	4.348	-17.150	-15.220	-12.499
2013	-11.657	6.886	-13.008	-11.364	-9.008
2014	-5.726	3.393	-7.460	-5.793	-4.058

Source: Own elaborations from Italian Credit Register data.

Table 2.4: Melitz–Polanec decomposition of Italian aggregate manufacturing productivity

	Avg. prod.	Reallocation	Entry	Exit	Total
2004	-0.21	2.89	-1.49	2.18	4.09
2005	-2.80	2.70	-1.52	2.00	0.51
2006	1.26	3.48	-1.81	2.18	5.82
2007	-0.39	4.18	-2.00	2.21	3.25
2008	-3.73	0.38	-1.94	1.89	-2.60
2009	-17.95	2.73	-1.48	2.91	-13.88
2010	1.02	7.69	-1.77	2.84	9.62
2011	-3.13	6.34	-1.89	2.40	2.50
2012	-7.52	2.88	-1.66	2.54	-4.12
2013	-5.60	6.72	-1.61	2.83	0.92
2014	-1.90	7.71	-1.71	3.17	6.03

Notes: Own elaborations from ASIA dataset. Productivity is measured as sales per worker. Aggregate productivity is defined as the weighted average of firm-level log productivities. The sum of the single components may not add up to the total variation, since (i) entry and exit from the manufacturing sector and (ii) false entry and false exits (Geurts and Van Biesebroeck, 2014) are not accounted for. Their combined impact on the dynamics of aggregate productivity is negligible relative to the one of the displayed components.

Table 2.5: Incumbents

	(1)	(2)	(3)
	(a) Productivity growth at the firm level		
CSS_{t-1}	0.0885**	0.00765	0.128**
	(0.0387)	(0.0410)	(0.0479)
Obs.	4,357,822	1,737,888	2,619,934
R^2	0.049	0.032	0.055
	(b) Firm-level contribution to reallocation		
CSS_{t-1}	-0.000000734*	-3.95e-08	-0.00000106**
	(0.000000378)	(0.000000613)	(0.000000426)
Obs.	4,357,719	1,735,912	2,621,807
R^2	0.035	0.033	0.037

Notes: All regressions include year×province and year×sector (2 digit) fixed effects. Firm level controls include size class and quintiles of the productivity distribution. Dependent variables and bank shocks expressed in percentage points. Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.6: Growth rate of employees in incumbent firms

	full sample	before crisis	after crisis
		Average effect	
CSS_{t-1}	0.0316*	-0.00816	0.0459*
	(0.0181)	(0.0291)	(0.0241)
		By productivity quintile	
$CSS_{t-1} \times Q1$	0.0680***	-0.0590	0.0720***
	(0.0190)	(0.0472)	(0.0231)
$CSS_{t-1} \times Q2$	0.0367**	0.00948	0.0456*
	(0.0174)	(0.0406)	(0.0225)
$CSS_{t-1} \times Q3$	0.0155	0.00803	0.0276
	(0.0202)	(0.0337)	(0.0258)
$CSS_{t-1} \times Q4$	0.0254	0.00175	0.0479*
	(0.0194)	(0.0301)	(0.0247)
$CSS_{t-1} \times Q5$	0.0175	-0.000887	0.0244
	(0.0210)	(0.0282)	(0.0275)
		By size class	
$CSS_{t-1}\times$ 0–1 employees	-0.0233	-0.0632*	-0.0226
	(0.0200)	(0.0364)	(0.0278)
$CSS_{t-1}\times$ 2–5 employees	0.0783***	-0.00409	0.102***
	(0.0198)	(0.0352)	(0.0250)
$CSS_{t-1}\times$ 6–9 employees	0.0389	0.0672	0.0653**
	(0.0293)	(0.0451)	(0.0315)
$CSS_{t-1}\times$ 10–19 employees	0.0257	0.0371	0.0405
	(0.0249)	(0.0338)	(0.0300)
$CSS_{t-1}\times$ 20–49 employees	0.0479**	0.0452	0.0667**
	(0.0220)	(0.0377)	(0.0287)
$CSS_{t-1}\times$ 50–100 employees	0.0296	0.0761	0.0355
	(0.0218)	(0.0479)	(0.0319)
$CSS_{t-1}\times$ 101–250 employees	0.0139	0.182***	0.0137
	(0.0211)	(0.0376)	(0.0295)
$CSS_{t-1}\times$ 250+ employees	-0.00280	0.256***	-0.00294
	(0.0230)	(0.0609)	(0.0264)
Obs.	4,296,940	1,699,338	2,597,602
R^2	0.014	0.013	0.013

Notes: All regressions include year × province and year × sector (2 digit) fixed effects. Firm-level controls include size class. Dependent variables and bank shocks expressed in percentage points. Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.7: Exiting firms

	(1)	(2)	(3)
	full sample	before crisis	during crisis
		(a) Exit rate	
CSS_{t-1}	-0.0586**	-0.0143	-0.0773**
	(0.0226)	(0.0326)	(0.0333)
Obs.	4,946,982	1,932,989	3,013,993
R^2	0.032	0.031	0.033
		(b) Relative size of exiting firms	
CSS_{t-1}	0.433***	0.335	0.472***
	(0.0965)	(0.218)	(0.116)
Obs.	353,439	132,017	221,422
R^2	0.114	0.109	0.118
		(c) Relative productivity of exiting firms	
CSS_{t-1}	-1.958***	-0.908*	-2.302***
	(0.331)	(0.445)	(0.355)
Obs.	353,366	132,667	220,699
R^2	0.092	0.094	0.091

Notes: All regressions include year × province and year × sector (2 digit) fixed effects. Regressions in panels (a) and (c) include size class as a firm-level control; those in panel (b) include quintiles of the productivity distribution, instead. Dependent variables and bank shocks expressed in percentage points. Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.8: Exiting firms in credit expansion vs shrinking times

	(1) log sales p.w.	(2) log employees	(3) Δ sales p.w.	(4) Δ employees
Exit in crunch years	0.0825***	0.00508	0.0340***	0.0314***
	(0.0116)	(0.0136)	(0.00222)	(0.00332)
Obs.	531,801	531,804	470,568	472,115
R^2	0.146	0.116	0.070	0.026

Notes: The sample of these regressions is a balanced panel of firms for which we can observe exit; for each firm, the observations refer to the three years prior to exit. All regressions include year×province and year×sector (2 digit) fixed effects. All regressions include a 2^{nd} order polynomial in firm age. Size class controls in columns (1) and (4); productivity class controls in columns (2) and (3). Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.9: Entrants

	(1)	(2)	(3)
	full sample	before crisis	during crisis
		(a) Entry rate	
CSS_{t-1}	0.0142	-0.0430	0.0336
	(0.0253)	(0.0456)	(0.0305)
Obs.	24,356	8,986	15,370
R^2	0.190	0.182	0.191
		(b) Relative size of entrants	
CSS_{t-1}	0.357***	0.343**	0.364**
	(0.0911)	(0.148)	(0.131)
Obs.	261,972	109,103	152,869
R^2	0.088	0.080	0.088
		(c) Relative productivity of entrants	
CSS_{t-1}	1.798***	1.322**	1.956***
	(0.260)	(0.496)	(0.344)
Obs.	262,024	108,626	153,398
R^2	0.088	0.077	0.093

Notes: All regressions include year × province and year × sector (2 digit) fixed effects. Regressions in panel (a) have been performed on data collapsed at the province × sector × year level. Regressions in panel (b) include quintiles of the productivity distribution as a firm-level control; those in panel (c) include size class, instead. Dependent variables and bank shocks expressed in percentage points. Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.10: Entrants in credit expansion vs shrinking times

	(1) log sales p.w.	(2) log employees	(3) Δ sales p.w.	(4) Δ employees	(5) dummy exit
Entry in crunch years	0.0213**	-0.00657	0.0362***	0.0654***	-0.0411***
	(0.00831)	(0.00953)	(0.00666)	(0.00622)	(0.00406)
Obs.	588,800	589,806	562,345	555,929	588,814
R^2	0.159	0.131	0.058	0.084	0.021

Notes: The sample of these regressions is a balanced panel of firms for which we can observe entry; for each firm, the observations refer to the three years after entry. All regressions include year×province and year×sector (2 digit) fixed effects. All regressions include a 2^{nd} order polynomial in firm age. Columns (1) to (4) also include a dummy for exit. Size class controls in columns (1), (4) and (5); productivity class controls in columns (2) and (3). Heteroskedasticity-robust standard errors clustered at the province and sector level in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

References

Abbate, C. C., M. G. Ladu, and A. Linarello (2017). An Integrated Dataset of Italian Firms: 2005-2014. *Bank of Italy Occasional Papers 384*, 3–26.

Amiti, M. and D. E. Weinstein (2013). How much do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data. NBER Working Paper No. 18890.

Berrospide, J. M. and R. M. Edge (2010). The Effects of Bank Capital on Lending: What do we Know, and what does it Mean? FEDS Working Paper No. 2010-44.

Bofondi, M., L. Carpinelli, and E. Sette (2017). Credit Supply during a Sovereign Debt Crisis. Forthcoming on *Journal of the European Economic Association*.

Buera, F. J. and Y. Shin (2013). Financial Frictions and the Persistence of History: A Quantitative Exploration. *Journal of Political Economy 121*(2), 221–272.

Caballero, R. J. and M. L. Hammour (1994). The Cleansing Effect of Recessions. *The American Economic Review 84*(5), 1350–1368.

Calligaris, S., M. Del Gatto, F. Hassan, G. I. P. Ottaviano, and F. Schivardi (2016). Italy's Productivity Conundrum. A Study on Resource Misallocation. European Economy Discussion Paper 030, European Commission.

di Patti, E. B. and E. Sette (2016). Did the Securitization Market Freeze affect Bank Lending during the Financial Crisis? Evidence from a Credit Register. *Journal of Financial Intermediation 25*, 54–76.

Foster, L., C. Grim, and J. Haltiwanger (2014). Reallocation in the Great Recession: Cleansing or Not? NBER Working Paper No. 20427.

Foster, L., J. Haltiwanger, and C. Syverson (2016). The slow growth of new plants: Learning about demand? *Economica 83*(329), 91–129.

Gamberoni, E., C. Giordano, and P. Lopez-García (2016). Capital and Labour (Mis)allocation in the Euro Area: some Stylized Facts and Determinants. *Bank of Italy Occasional Papers 349*, 3–57.

Geurts, K. and J. Van Biesebroeck (2014). Job Creation, Firm Creation, and *de novo* Entry. CEPR Discussion Paper No. DP10118.

Gopinath, G., S. Kalemli-Ozcan, L. Karabarbounis, and C. Villegas-Sanchez (2017). Capital Allocation and Productivity in South Europe. Forthcoming on *Quarterly Journal of Economics*.

Greenstone, M., A. Mas, and H.-L. Nguyen (2014). Do Credit Market Shocks Affect the Real Economy? Quasi-experimental Evidence from the Great Recession and "Normal" Economic Times. National Bureau of Economic Research.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Iyer, R., J.-L. Peydró, S. da Rocha-Lopes, and A. Schoar (2014). Interbank Liquidity Crunch and the Firm Credit Crunch: Evidence from the 2007–2009 Crisis. *Review of Financial studies 27*(1), 347–372.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina Salas (2010). Credit Supply: Identifying Balance-sheet Channels with Loan Applications and Granted Loans. ECB Working Paper No. 1179.

Khwaja, A. I. and A. Mian (2008). Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market. *The American Economic Review 98*(4), 1413–1442.

Larrain, M. and S. Stumpner (2012). Understanding Misallocation: The Importance of Financial Constraints. Columbia University Working Paper.

Lee, Y. and T. Mukoyama (2015). Entry and Exit of Manufacturing Plants over the Business Cycle. *European Economic Review 77*, 20–27.

Manaresi, F. and N. Pierri (2016). Credit Constraints and Firm Productivity: Evidence from Italy. Mimeo. Available at `https://ssrn.com/abstract=2906809`.

Melitz, M. J. and S. Polanec (2015). Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit. *The RAND Journal of Economics 46*(2), 362–375.

Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *The American Economic Review 104*(2), 422–458.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *The American Economic Review 104*(10), 3186–3221.

Olley, G. S. and A. Pakes (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica 64*(6), 1263–1297.

Schivardi, F. (2003). Reallocation and Learning over the Business Cycle. *European Economic Review 47*(1), 95–111.

Schivardi, F., E. Sette, G. Tabellini, et al. (2017). Credit Misallocation During the European Financial Crisis. CEPR Discussion Paper No. DP11901.

Chapter 3

Mechanisms of the Urban Productivity Premium: Evidence from Italian Firms

Joint with Andrea Lamorgese, Bank of Italy.

3.1 Introduction

Urbanization is not a novel concept in the history of mankind. Economic development and the enhancement of living standards have often been associated with the formation and the expansion of cities. During the XX century urban population grew above the national average in many countries (Table 3.1), and contemporary times make no exception, with the emerging economies experiencing the transition of huge masses of people from rural areas to the cities. The United Nations estimate that in 2007 —for the first time in human history— people living in urban areas have surpassed in number those that live in the countryside (UN, 2013). As a result, economic activity got progressively more concentrated in granular entities across space, that therefore contribute increasingly more to global growth. In 2010, 85% of US GDP was generated in large cities; the same statistics was 78% in China, 76% in Latin America and just under 65% for the Western European countries as a whole (Manyika et al., 2012). Italy, which will be the object of this study, has one third of its value added produced by the 12 largest urban areas, while another third is produced in the following 67 ones.

The success of cities as a form of social and economic organization has led scholars to investigate the driving forces behind it. One of the more soundly-established features of cities is that they display a higher productivity with respect to rural areas, and that such advantage tends to increase with city size.[1] Ever since Marshall (1890), economic literature has especially emphasized the role played by agglomeration economies in shaping such a productivity premium. This definition generally groups all positive externalities that may arise in an environment characterized by a high density of agents (either individuals or firms); labor pooling, cost sharing and knowledge spillovers are generally regarded as the main channels through which agglomeration economies operate (Glaeser, 1999; Duranton and Puga, 2004). Besides agglomeration, additional explanations for the observed urban productivity advantage have been put forth: one of them relates to the endowment of exogenously-determined factors, such as natural resources or other amenities; another points at the potential sorting of more productive firms and/or more skilled workers to urban areas; finally, stronger selection mechanisms could be at work in more competitive environments such as cities, thus leading only the most productive firms to survive.[2]

The increased availability of micro-data has sustained the flourishing of studies that aim at quantifying the relative importance of these alternative explanations. Most of this empirical body of work has focussed on wages —the flip side of productivity in equilibrium— to explore these issues: making extensive use of matched employer-employees data, they have studied the mechanisms that sustain the urban wage premium, showing that people with higher ability tend to sort to cities (Yankow, 2006; Combes et al., 2008; Eeckhout et al., 2014),[3] that the assortative matching between more productive firms and more skilful workers is stronger in urban areas (Andersson et al., 2007; Dauth et al., 2016), that agglomeration economies are more relevant for non-routine jobs (Andersson et al., 2014), that the workers may accrue a wage premium both instantaneously (D'Costa and Overman, 2014) and through a faster wage growth over time (Glaeser and Maré, 2001), and that the experience accumulated in cities is more valuable and persistent (Baum-Snow and Pavan, 2011; De La Roca and Puga, 2017).

In this paper we adopt the perspective of the firm and provide a direct ob-

[1]Ahrend et al. (2014) provide cross-country evidence. In Figure 3.1 we verify this stylized fact for Italy as well: panel (a) shows that average labor productivity scales up with the size of the Local Labour Market (LLM); panel (b) shows that urban areas (and especially the big ones) consistently display higher levels of labor productivity.

[2]Behrens and Robert-Nicoud (2015) offer an extensive discussion of the various sources of urban productivity premium, providing some related stylized fact for US cities.

[3]Glaeser et al. (2001) show that the demand for living in cities (as captured by rents) has grown faster than earnings, thus suggesting that there are other reasons besides pure labor market considerations that drive people —ad especially more skilled ones— to urban areas.

servation of the urban productivity premium using detailed firm-level data on the universe of non-agricultural and non-financial Italian firms. In the first part of the paper, we investigate the sources of this premium using the toolbox of the urban wage premium literature. More specifically, we adopt an empirical approach similar to the one in Abowd et al. (1999) to disentangle firm- from city-specific components of firm labor productivity (measured as value added per worker). The separate identification of these two effects is guaranteed by the fact that we are able to observe firms moving across cities over time; this is a feature of our data that we will extensively exploit in this paper. We conceptually regard the firm-specific effects as capturing the sorting component (that in principle may be either positive or negative) of the observed urban productivity premium. On the other hand, throughout most of the paper we will be thinking at the city-specific component as capturing the extent of agglomeration economies within each urban area. Nonetheless, our framework is not able to discriminate between "pure" agglomeration economies and other potential sources of urban advantage, such as locational fundamentals (amenities) or harsher firm selection. Using the same data, Lamorgese and Petrella (2016) show that the higher education level of the workforce and the presence of thicker labor markets are among the more relevant local characteristics associated with the urban productivity premium, while competition seem to play a minor role. Besides confirming some of the results of the urban wage premium literature, these findings suggest that traditional sources of agglomeration economies are predominantly at work in Italian cities.

In the second part of the paper, we focus on relocating (or switching) firms as an interesting analytical object to unveil some of the mechanisms through which firms appropriate of the productivity advantage offered by a urban environment. Since it allows us to observe the same firm in two distinct cities, we consider firm relocation as an ideal experiment to evaluate the advantages of locating in a urban areas, provided that we are able to find adequate counterfactuals that instruct us on the evolution of firm productivity, hadn't the relocation occurred. The aim of this exercise is twofold. On one side, we want to assess whether a firm gains in productivity when it relocates its activities, and whether such gain is higher for firms switching to urban areas; on the other side, we want to study how these productivity gains accrue over time, with the aim of understanding whether operating in a urban environment is a static advantage —that accumulates once and for all in the year of switch— or if it instead materializes only through time (that is, upon learning). We first document some notable facts on relocating firms, namely that they typically are larger, more productive and younger than static firms. We then estimate the gains from relocation by means of a standard difference-in-differences approach, in which we control for firm-specific time-invariant potential confounders to identify the effect of treatment (that is, of relocation). This empirical strategy crucially relies on the assumption of parallel trends between the group of treated and control firms; this is however difficult to sustain, precisely because of the documented features of switching firms. As a consequence, we adopt a synthetic control strategy, that builds on the works by Abadie and Gardeazabal (2003) and Abadie et al. (2010), to build artificial counterfactuals that share with the switching firms the pre-treatment behaviour of a number of relevant variables. The peculiarity of this exercise resides in the fact that we have multiple treated units (and synthetic controls), which calls for a re-aggregation of the estimated effects. We exploit the fact that firms switch at different points in time to net the relocation gain from year-specific components common to all switching firms.

Our results indicate that the intrinsic qualities matter more than the environment in which the firm operates: the firm-specific component explains a substantial fraction of the variability in firm-level productivity, while the share explained by the city-specific components is significantly lower, despite being non-negligible. Sorting and agglomeration economies are both at work in shaping the productivity premium of urban areas: the elasticity of both firm- and city-specific components with respect to city size is positive, meaning on one hand that firms with higher intrinsic qualities sort into bigger cities, and on the other that agglomeration economies scale up with the size of urban areas; the elasticity of the former is however more sizable than that of the latter. The difference-in-differences and the synthetic control exercises unveil significant rewards from switching location: on average, a relocating firm experiences a productivity gain of roughly 10 percentage points at a 4-years distance from the switch; the gain is positive irrespective of the city of destination, but it is highest for firms relocating from non-urban to urban areas, while it is lowest for firms going the other way round. Analyzing the patterns of productivity accrual provides further insights on the nature of urban agglomeration economies. A (minor) part of the urban productivity premium has a static nature, meaning that firms moving to a urban environment immediately accrue it in the year of the switch; such gain appear to be transitory, since firms moving away from cities specularly suffer a productivity drop. This mechanism is stronger for bigger urban areas. Most of the relocation gains, however, accrue over time, irrespective of the destination of the switch; the rate of accumulation, though, is higher for firms that move from non-urban to urban locations, hinting at faster learning processes for this category of firms. This is also consistent with a higher reward for an extra year of experience spent in a urban environment, and with the entity of such reward being decreasing in the amount of accumulated urban experience (that is, gains are higher for firms that come from non-urban environments and therefore have no urban experience). Additional evidence supportive of these mechanisms is also provided in a regression framework *à la* De La Roca and Puga (2017).

We are not the first ones to study the productivity premium for a firm to operate in a urban environment. Henderson (2003) and Moretti (2004) estimate plant-level production functions to study extent of different externalities on the performance of firms; while the former focuses on scale externalities from other plants within the same industry and from the size of local economic activity outside the industry, the latter assesses the existence of human capital spillovers, showing that the productivity of plants is affected by the availability of skills on the labour market. Combes et al. (2012), instead, lay down a model that entails both selection and agglomeration mechanisms; they test the predictions of the model against establishment-level data, showing that selection alone cannot explain the observed urban productivity premium. Gaubert (2017) lays down a model to study the sorting of firms across space and structurally estimates it using French firm-level data; she find that two thirds of the productivity premium of larger cities can be attributed to sorting. Di Giacinto et al. (2014) also use Italian data to show that cities display a higher productivity premium with respect to industrial districts; they attribute this pattern to stronger agglomeration economies operating in urban environments, only partly attributable to the different skill composition of the workforce. While contributing to this literature, our paper also draws a parallel with the literature on urban wage premium, from which it borrows the analytical concepts and the empirical methods. In this perspective, this paper is the first to our knowledge that is able to provide a direct quantification of the agglomeration economies at the city level, whose identification is only made possible by the fact that we observe firm switching location.

This paper also sheds some light on the characteristics of switching firms and on the gains connected with a relocation episode. Empirical evidence on firm relocation is scant. Brouwer et al. (2004) analyze cross-country data on large firms (more than 200 employees) to investigate the determinants of the firm relocation decision, finding that firms involved in mergers and acquisitions or serving larger markets have a higher propensity of relocating. Knoben et al. (2008) use survey data collected among managers of Dutch automation services firms to investigate the effect of relocation, finding that moving to urban areas harms the firm's performance. Both of these studies use highly selected samples, and therefore it's difficult to draw a parallel with our results. Bergeaud and Ray (2017) write down a model of the relocation decision and estimate it using French firm-level data, finding that most of the relocations involve moves within a short distance from the departure point, and that higher relocation costs (that is, higher prices in the real-estate markets) lowers the firms' propensity to move, thus hindering the job creation of more productive firms. The broad stylized facts on relocating firms presented by Bergeaud and Ray (2017) are coherent with many of the features characterizing the switching firms in our data. Our paper adds to this literature in two ways: first, it documents some basic stylized facts on relocating firms; second, it provides an estimate of the productivity gains connected with relocation, relying on an extensive dataset covering the universe of Italian firms and testing the results against different identification strategies.

The remainder of this paper is organized as follows. Section 3.2 introduces the definition of urban area that we will use throughout the paper. Section 3.3 presents the firm-level data used for our empirical analyses. Section 3.4 studies the urban productivity premium, disentangling firm- and city-specific characteristics, and giving an assessment of the relative contribution of sorting and agglomeration economies in shaping the urban productivity premium. Section 3.5 focuses on relocating firms, describes the difference-in-difference and the synthetic controls exercises, and analyzes the pattern of productivity accrual following a relocation episode. Section 3.6 concludes.

3.2 Urban area definition

Before delving into the description of the data used in this paper, it's worth spending a few words on the definition of city, which is not a univocal economic concept in the empirical urban literature. Indeed, urban areas do not necessarily overlap with the administrative borders of single municipalities, and in certain cases not even with those of the administrative units at a lower level of breakdown (NUTS3 regions, provinces in Italy). As a consequence, urban areas end up being constituted by more than one municipality, but often their territory does not coincide with the one of a province. In a word, the city is a level of territorial aggregation that lies in between the municipality and the province, but that does not share with them an independent administrative logic, since it is a rough aggregation of distinct administrative units.

A functional agglomeration which lies in between municipalities and provinces is the Local Labor Market (LLM), which is conventionally deemed as a good representation of a spatial agglomeration. LLMs are the result of a partition of the national territory, obtained by aggregating municipalities in such a way that they contain both the place of residence and the workplace of (a majority of) individuals.[4] The definition of urban areas used in this paper borrows from both the map of the Italian LLMs in 2011 and the OECD–Eurostat classification of cities; the latter is used to distinguish the LLMs that have urban nature from those that do not.

The OECD–Eurostat classification singles out cities according to a density-based criterion. The underlying methodology defines a urban area as a homogeneous set of territories whose population density exceeds a certain threshold (1,500 inhabitants per square km). This definition is consistent with the traditional view that urban agglomerations are the places where production and knowledge spillovers take place, for the simple reason that density creates thick markets and favours the matching between demand and supply. It is therefore consistent with the traditional sources of agglomeration (labor pooling, cost sharing and knowledge spillovers).[5]

In this paper we will label as urban areas all the LLMs that contain at least one city according to the the OECD–Eurostat classification; the LLMs that do not contain any city are labeled as non-urban. Within the group of urban areas, we further distinguish into big and small ones, according to a population threshold of 500 thousand inhabitants.[6] Over the 611 Italian LLMs in the 2011 classification, we identify 73 urban areas (or urban LLMs); six of them are big (over 500 thousand inhabitants), and they correspond to the LLMs of Roma, Milano, Napoli, Torino, Palermo and Genova. Figure 3.2 shows the geographical distribution of the urban areas, which tend to be concentrated in regions having higher levels of economic activity (in Northern Italy) or hosting big industrial factories (like in certain regions of the South); urban areas are less frequent in the regions of the Center, also because of the more mountainous nature of the terrain. Over the past three decades, the 73 urban areas have grown both in population and size, having absorbed an increasing number of municipalities: as Table 3.2 shows, non-urban areas have diminished over time from 880 to 538 over the 4 Censuses between 1981

[4]At the end of 2014, the Italian statistical agency (Istat) issued the fourth classification of LLMs, based on the commuting flows detected in the 2011 wave of the Census (see `http://www.istat.it/it/strumenti/territorio-e-cartografia/sistemi-locali-del-lavoro`). The three previous classifications —based on the commuting flows of the 1981, 1991, and 2001 waves of the Census— used the same kind of data but slightly different definitions. Starting in 2011, the definition has changed to be consistent with the European definition of LLM; for the sake of comparability, new and old definitions coexist for year 2001.

[5]A detailed description of the OECD-Eurostat methodology can be found in Dijkstra and Poelman (2012).

[6]This threshold is often adopted by national statistical agencies to single out the metropolitan areas.

and 2011.[7].

3.3 Data

The main data source used in this paper is the result of a joint collaboration between the Bank of Italy and the Italian National Statistical Agency (ISTAT), obtained by combining the information of the Italian Register of Active Firms (ASIA) with data retrieved from other statistical, administrative and fiscal sources. The dataset contains firm-level information on the universe of non-agricultural and non-financial Italian firms in the years 2005–14. For each firm, the dataset gathers information on the incorporation year, the industry in which the firm operates (at the 4-digit breakdown of the Nace rev. 2 classification), the number of employees, the sales, and the value added. Data on sales and value added are deflated to 2010 prices, using sector-level price indexes for sales and value added, respectively.[8] Most importantly, the dataset contains information on the municipality in which the firms are located in each year. The fact that we are able to observe the position of firms in time allows us to single out relocation episodes, that we will exploit throughout the paper either to extract the city-specific component of the urban productivity premium, or to study the process of productivity accrual connected with a relocation event.

Data on value added are retrieved using a variety of administrative and fiscal sources. For the years before 2012, some imputations were needed in order to recover the value added for all firms.[9] Starting from 2012, the database corresponds to the data of the FRAME-SBS archives, which are also the microeconomic information at the base of the national accounts. In the aggregate, our data are consistent with Italian national accounts and with the Structural business statistics (SBS) from Eurostat.

Table 3.3 shows some descriptive statistics based on our dataset, distinguishing for firms located in urban and non-urban areas. The number of firms in the whole economy has been rising before the crisis (up to 2008)

[7]Further stylized facts on Italian urban areas can be found in Lamorgese and Petrella (2017)

[8]Deflators vary at the 2-digit Nace rev. 2 disaggregation. See Abbate et al. (2017) for a detailed description of the dataset.

[9]With the existing data, it is possible to recover value added information for a share of firms varying between the 80% and the 95% of the universe of firms in the years 2005–2011. For the remaining firms, the value added is imputed according to the median value added per employee in the same region, size class and industry. Imputation of missing values is called for in order to be consistent with aggregate national accounts data. Dropping the firms with imputed values doesn't change the estimates contained in this paper.

and steadily decreasing afterward, both in urban and non-urban areas. This dynamics has been primarily driven by manufacturing, where a widespread process of restructuring, churning, and selection had been going on since the early Nineties; the number of firms in services has shown, instead, much smaller fluctuations over the period of observation.

Average firm size —measured as average employees per firm— has shown roughly the same hump-shaped trajectory, although remaining in every year larger for urban firms than for non urban ones. Labor productivity, that we measure as value added per worker throughout the paper, has been slightly declining in both urban and non-urban areas during the period of observation, though with some fluctuation.[10] Throughout all years, labor productivity is higher by about 30% for urban firms: this is a first and very rough evidence of the urban productivity premium that we will more soundly establish in the next section and that will be the main focus of this paper.

3.4 Disentangling sorting and agglomeration components of labor productivity

In this section we first aim at assessing the existence of a urban productivity premium, and then at studying what are the driving forces behind it, focussing on two main channels. The first one, that we will label as sorting or positive assortative matching, relates to the fact that more productive firms may self-select into bigger and more urbanized economic environments. The second is specific to each city, and is conceptually linked to the agglomeration economies that arise in denser areas, traditionally regarded as the outcome of thicker markets, favouring the efficient matching between demand and supply of labor, the interchange of ideas and the diffusion of innovation.

To do that, we adopt an approach similar to the one that Abowd et al. (1999) use on matched employer-employee data to disentangle worker- from firm-specific effects in a wage equation. We will structure our estimation strategy in two steps: in the first place, we will extract the city and firm characteristics in a regression on firm-level log labor productivity; subsequently, we will assess if and to what extent these characteristics scale with the size and the urban nature of the city.

[10]The decline in labor productivity has not interested all sectors. In the period of observation labor productivity has been mildly increasing in manufacturing —also as a consequence of an increased role for reallocation and selection— and decreasing in services (Linarello and Petrella, 2017).

Therefore, we will start by estimating the following equation:

$$\pi_{iy} = X'_{iy}\beta + \delta_c \times \delta_y + \delta_f + \varepsilon_{iy} \tag{3.1}$$

where the dependent variable π_{it} is the log VA per worker of firm i in year y; X_{iy} is a matrix containing a set of —possibly time-varying— firm-level controls, including age (linear and squared), size class and industry (4-digit Nace rev. 2) fixed effects; δ_y are instead year fixed effects. δ_c and δ_f are a battery of city and firm fixed effects, respectively, whose estimation represent the objective of this first step of the analysis. The first ones capture the city-related component of productivity, after controlling for firm's characteristics; as such, it gauges the extent of the agglomeration economies at the city-level. We have interacted them with the year fixed effects, in order to let the extent of the agglomeration economies vary through time, thus providing more flexibility to the model.[11] Firm fixed effects, instead, capture the intrinsic quality of a firm, thus capturing the sorting behaviour of more productive firms in certain areas; we keep them fixed in time.[12]

The identification of these effects is possible due to the fact that we observe firms relocating across cities; although the fraction of relocating firms is relatively small (the 7% of the firms in our sample), the relevant fact for the identification of our effects is that the wide majority of the observed firms and cities (more than 99% of them) belong to a single large connected group.[13]

For a clean identification of the firm and city fixed effects, our empirical strategy also requires the additional assumption of random mobility of firms across cities, conditional on observables.[14] Despite being a demanding claim when dealing with firm mobility, we will argue that it is a reasonable

[11] We have also estimated equation 3.1 keeping the city component fixed in time. Since adopting this alternative specification does not imply any significant change in the following analyses, we most of the times stick to the more flexible formulation in equation 3.1.

[12] Macis and Schivardi (2016) —applying a similar framework to disentangle the firm and worker components of wage— discuss the implications of letting both dimensions vary through time, and show how to implement it. In the context of our exercise, however, we chose to keep the firm effect fixed for two reasons: on one hand, we don't have a prior on potential asymmetrical shock that may have differentially affected certain firm categories (Macis and Schivardi (2016) argued that the 1992 devaluation of the lira had changed the relative market value of skills); on the other hand, we observe relatively few years, and the intrinsic quality of a firm is unlikely to vary over short time windows.

[13] In our setting, a connected group is defined such that it contains (i) all the firms that have ever been located in any city of the group, and (ii) all the cities in which any firm of the group was ever located (Abowd et al., 2002).

[14] This parallels the assumption of random mobility of workers across firms discussed in Abowd et al. (1999).

assumption for our sample of firms. It the first place, one might be worried that the relocation of firms may be driven by idiosyncratic city-level shocks, implying a migration of firms to cities receiving a good shock in a certain year. However, it is unlikely that the benefit reaped from taking advantage of a short term shock overcomes the significant sunk costs that the firm has to sustain for the relocation. This point is further addressed by the fact that equation 3.1 entails year-varying city fixed effects: by construction, this extracts the city-year component of firm productivity, and therefore removes any concern of correlation between firm mobility and economic cycle at the city level. A second concern regarding the identification of the fixed effects is related to the fact that firms with positive transitory productivity shocks may sort to cities with higher agglomeration effects. We argue that this is not the major driver of firm mobility in our sample, based on two pieces of evidence that we provide with more details at the beginning of next section. First, relocating firms do not display particular pre-relocation productivity trends (conditional on observables), and this fact holds even if we allow for the effects to be heterogeneous across destinations. Second, we show that most of the observed firm movements occur within a short distance from the departure point; as a consequence, we claim that the relevant driver for firm mobility is most of the times the necessity of expanding the firm's premises, rather than the response to a positive shock or the quest for a match with a better-performing city.

Before turning to the estimation of equation 3.1, we first assess the existence of a urban productivity premium in Italian cities. In Table 3.4 we regress the firm-level log value added per worker against the log population of the LLM the firm belongs to. In column (1) we only control for year fixed effects, while in column (2) we add firm-level controls such as age (linear and squared), size and sector fixed effects. The correlation is positive and strongly significant, meaning that firms in larger LLMs display higher labor productivity. As expected, in all specifications the return to age is positive, though at decreasing rates (as captured by the curvature term, represented by age squared). In columns (3) to (6) we then directly address the issue of urban areas, by substituting the log population with dummies that capture the urban nature of the LLM the firm belongs to; in every case, the excluded category is represented by the firms located in non-urban areas. The productivity premium of urban areas is conspicuous: net of observable characteristics, firms in located in a city are on average 9.3% more productive than their non-urban counterparts. In the case of big urban areas, this premium jumps up to 12.1%.

As discussed above, estimating equation 3.1 will help us assessing the relative importance of two (potentially large) sources of the urban productivity premium we have just documented: on one hand, the productivity advantage may be just the result of the fact that intrinsically better firms sort into urban environments; on another hand, it could be that the urban nature of the city endows firms with a greater scope for productivity enhancement. Table 3.5 contains the correlations between the firm-level log productivity (π) and the city and firm fixed effects (δ_f and δ_c); we report the results for both the time-varying and the fixed city-component specifications. The results show that both forces are at work, although with different intensities. The role of sorting is prevalent, since the intrinsic characteristics of the firms explain a relevant fraction of the observed variation in firm-level productivity (correlation of 0.8 between π and δ_f). The weight of agglomeration economies —as measured by the correlation between productivity and the city component (0.16)— is much smaller but non-negligible. As expected, the correlation between firm and city components is close to zero, since they are by construction orthogonal. Moreover, the city and firm fixed effects estimated under the two different specifications display a strong positive correlation.

Another test of these patterns is offered in Tables 3.6 and 3.7, where we regress the city and firm effects against the log population or the urban nature of the corresponding LLM. As expected, both effects scale with the size of the urban area, meaning that bigger cities both attract intrinsically more productive firms and provide agglomerative advantages to the firms operating there; the size of the effect is nonetheless bigger for firm effects. The same patterns are confirmed when considering urban versus non-urban areas, as in columns (2) and (3).

Having assessed the relative importance of sorting and agglomeration in shaping the size of the urban productivity premium, in the next section we rule out sorting and study whether the advantages deriving from agglomeration are instantaneously appropriated by firms, or if instead they are accrued through time in a sort of learning process. To do that, we exploit the features of a peculiar group of firms that relocate their activities across space during our observation period.

3.5 The productivity gains of relocation

As argued in section 3.3, one of the most interesting features in our dataset is that we are able to observe firms moving from one municipality to another. In section 3.4 this feature has allowed us to extract the city-specific component of the urban productivity premium under the framework of Abowd et al. (1999). In this section, we take a closer look to the firms that switch the municipality in which they are located (henceforth called switchers, movers or relocating firms); we will first show that these firms are peculiar in many respects, and we will then argue that they represent an interesting analytical object to shed some light on the mechanisms that govern the accrual of productivity gains in different urban environments. While in section 3.4 we have disentangled the role played by the sorting of better firms to larger cities from the one played by agglomeration economies, it is still unclear whether the nature of these agglomeration economies is static or dynamic (i.e. if the gain accrues instantaneously or over time, in a kind of learning process). We argue that relocating firms can help us improving our knowledge on this topic: since we observe the same firm in two distinct urban environments over a satisfactorily long time span, the productivity gain of relocation could in principle be measured by simple difference, if only we were able to find a proper counterfactual for the firm's productivity, hadn't the switch happened.

Finding a counterfactual for moving firms is by no means a trivial task, since movers display stark differences with respect to static firms. In our sample we observe 7,492,067 firms over an average time span of 5.8 years; 522,215 of them (the 7%) switch location during the period of observation. Most of relocating firms move only once, but 12% of them move twice or more. Relocations do not generally involve long-range displacements: as depicted in Figure 3.3, half of the moves are realized within 14 km from the departure point, while the 75th percentile of the distances distribution is equal to 35 km. Table 3.8 shows some basic facts on the transition probability across different types of areas, distinguishing between non-urban, small urban (those below 500 thousand inhabitants) and big urban areas (those above 500 thousand inhabitants). As a consequence of the low incidence of long-distance relocations, most of the switches (around 70%) occur within the same type of area, with no dramatic differences across departure points; as a matter of fact, 53% of the moves are realized within the same Local Labour Market. While a firm in a big urban area has roughly the same probability of "downsizing" to either a non-urban or a small urban area, the firms in non-urban or small urban environments find it more difficult to relocate to a big city, probably because of the higher sunk costs connected with such a move (higher salaries and real-estate costs).

Table 3.9 displays some descriptive statistics associated with relocating firms before the switch takes place: they are significantly larger (both in terms of employees and of sales), more productive, and younger than the firms that do not move during our observation period.[15] The productivity

[15] Remember that our dataset contains the universe of Italian firms; hence, it also includes very small and unipersonal firms, that rarely relocate.

premium for switchers survives even if we condition on size, sector, legal form, year and age, as in column (1) of Table 3.10: even before any relocation takes place, switchers-to-be are 8% more productive (in terms of VA per worker) than static firms. Column (2) looks at how the productivity premium for switchers evolves in the years preceding a relocation episode: the premium maintains roughly constant throughout all the years, and —if anything— slightly reduces in the three years before relocation, as if the firms were diverting some resources to prepare the future move. Columns (3) and (4) investigates the possibility that the switchers may be *ex-ante* systematically different, depending on the kind of relocation that they are about to undertake: the most productive firms are those that move from a urban area to another, while the least productive ones (though still being more productive than stayers) are those that move from one non-urban area to another one. Column (4) adds the size of the urban area to the picture: as a general rule, firms that "upgrade" —either by moving *ex-novo* to a urban area or by relocating to a bigger urban area— tend to be more productive than firms that "downgrade" to smaller cities or rural areas.

These facts provide us with a rather rich picture of relocating firms. They generally are young and productive firms that probably switch location either to further expand their activity levels (e.g. by acquiring bigger premises; this could especially be the case for short-distance moves) or to reap the benefits of thicker markets or agglomeration economies. In the remainder of this section we will investigate whether relocation grants an additional productivity gain to these firms, if this gain differs according to the origin or the destination of the firm, and if it accrues instantaneously or over time.

3.5.1 A difference-in-differences approach

For what we have discussed so far, it is clear that relocating firms are very peculiar objects. Exactly identifying the productivity gain associated to moving to a certain location is therefore a challenging undertaking, since in principle switching firms may differ from stayers along many dimensions than we are not able to control for, and that may in turn be correlated to our outcome variable and to the type of relocation that takes place (e.g. to a big/small urban area vs a non-urban one). To tackle this problem in a regression framework, we will first exploit the richness of our data, by setting up a Difference-in-Differences strategy, in which we consider a firm's relocation as the treatment of interest. The basic idea is to control for time-invariant potential confounders through the inclusion of firm fixed-effects; we will therefore be exploiting the time variability of our dataset to identify the effect of a switch. In formula:

$$\pi_{iy} = \alpha + \beta S_{iy} + \delta_i + \delta_y + \varepsilon_{iy} \tag{3.2}$$

where π_{iy} is log VA per worker of firm i in year y, S_{iy} is a dummy equal to one from the year in which the switch takes place onwards, δ_i and δ_y are firm and year fixed effects, respectively, and ε_{iy} is an error term. In order to be clear and consistent regarding the time in which the switch takes place, firms relocating multiple times have been excluded from the regression. Our coefficient of interest is β, which gauges the percentage productivity gain/loss arising as a consequence of a relocation.

In other specifications we check whether the overall effect captured by β has some margin of heterogeneity according to the type of switch that takes place, by substituting S_{iy} with a multiple-treatment categorical variable that basically captures the combination of the departure area with the destination one, in the spirit of the transitions presented in Table 3.8. We adopt different levels of detail in these kind of regressions, sometimes distinguishing —within urban areas— between big and small cities (based on the 500 thousand inhabitants threshold).

Finally, we look at the time profile of the productivity gain associated with a relocation, interacting the treatment with a set of dummies that capture how many years have passed from the relocation episode. In other words, we will be estimating the following:

$$\pi_{iy} = \alpha + \sum_{t=0}^{8} \beta_t S_{iy} \times I_{iy}^t + \delta_i + \delta_y + \varepsilon_{iy} \tag{3.3}$$

where I_{iy}^t is equal to one if firm i in year y is t years away from a switch. The summation runs from 0 —which is the value at which we we conventionally set t in the year of the relocation— to eight —which is the maximum number of observable periods, given the time span covered by our data. Also under this specification, we replace the treatment variable with a set of dummies that capture —with different levels of detail— the nature of the switch, based on the departure-arrival information. The intention is again that of highlighting potential margins of differentiation for the average effect estimated in equation 3.2, either across time or across treatment types.

Table 3.11 contains the estimates for the model in equation 3.2. Being the result of an independent choice of supposedly rational agents, we expect the relocation to have on average a positive effect, irrespective of the type of transition performed by the firm. The results in the table confirm that this is the case, with an average productivity gain of 6.8 percentage points with respect to static firms. This effect is heterogeneous across transition types, being bounded between the 10 percentage points of the firms switching from a non-urban to a urban area and the 4.8 percentage points for the firms moving the other way round. The same ranking holds in we distinguish between big and small urban areas: productivity premia are bigger for firms “upgrading” to urban or bigger urban areas, and smaller for “downgrading” firms.

Figures 3.4 and 3.5 graphically present the results arising from the estimation of equation 3.3, where we allow the effect to vary in time after the relocation takes place.[16] In the year of the relocation, the productivity gains are generally estimated to be positive, though being bigger for firms going to urban areas. The evolution profile over time shows that firms moving to urban areas also show a much faster growth of the productivity gain, that reaches nearly 20 percentage points at a 8-years distance from relocation; the productivity profile associated with the other types of transitions is less steep, touching 13 percentage points over the same horizon. Adding more details on the size of urban areas, such as in Figure 3.5, only makes the gains associated with moving to a big urban area more pronounced (both on impact and over time), while showing roughly the same picture for firms downgrading to smaller or non-urban areas. Overall, these results point to a higher productivity premium associated with (big) urban areas, along with a larger room for the dynamic accrual of productivity (learning channel) in urban environments.

The validity of the difference-in-differences strategy crucially relies on the assumption that —hadn’t the treatment occurred— treated and non-treated firms would have followed parallel trends over time. This would ensure that —net of the individual fixed effects— the differences registered after the treatment could be given a direct causal interpretation. In our case, however, this assumption is likely to be flawed, since we have already shown that relocating firms are profoundly different from static ones, even before any treatment (that is, a move) takes place; under these conditions, it is very difficult to defend the assumption of parallel trends. To provide more credible estimates of the effect of a firm relocation, we therefore depart from a pure regression approach and exploit the richness of our data by implementing a synthetic control strategy.

3.5.2 Synthetic controls

In this section we adopt the approach of Abadie and Gardeazabal (2003) and Abadie et al. (2010) as a method to find plausible counterfactuals for our relocating firms. The basic idea is to build an artificial (“synthetic”) firm

[16]The quantitative results are reported in Tables 3.15 and 3.16 of the appendix.

as a weighted average of other existing firms ("donors"); weights are determined so that they minimize the distance with the treated unit along a set of relevant characteristics (which include the outcome variable) in the pre-treatment period. In this way, we force the balancing of the pre-treatment characteristics between the treated firm and its synthetic control, thus allowing us to claim that the registered discrepancies after the treatment are entirely attributable to the relocation. Compared to other techniques —such as propensity score matching— used to circumscribe a control group, this approach not only performs a match based on a set of arbitrary observable covariates, but also aims at minimizing the pre-treatment distance of the outcome variable, in an effort of capturing potential unobservable confounders that may differentially affect the trends of treated and controls. In other words, this is a method to forcibly restore the parallel trend hypothesis, provided that we can find a sufficiently good linear combination of donors to approximate the pre-treatment behaviour. To this effect, the data we have in hand are helpful, since they provide us with an immense pool of potential donors.

The nature of our data entail that we do not have just a single treated unit, but we observe multiple treatments in our data (that is, multiple firms switching location). As a consequence, we build a synthetic control for each of the treated firms and then re-aggregate the results to provide a quantification of the average effect of relocating to another city. Instead of representing a difficulty, this richness of treatments is useful for netting out the common year-specific component of the productivity gain: if we think that the relocation premium may vary according to the realized general economic conditions, pooling together firms moving in different years gives us a mean to extract an effect which is independent of the year in which the move takes place. Given the number of years spanned by our dataset, our synthetic control analysis will only be focussing on the firms relocating in three years, from 2010 to 2012. This choice is motivated by the need of having a sufficiently long stretch of years before the treatment to compute the aggregation weights for the donors, while at the same time remaining with enough years left after the relocation to evaluate its effects.

As a preliminary step in our search for proper donors and weights, we first restrict the sample of potential donors according to certain observable characteristics. First of all, we do not include other switching firms in the group of donors. In the second place, we choose the donors so that they belong to the same province, kind of LLM (non-urban, small urban or big urban) and sector (3-digit Nace rev. 2 classification) as the treated unit. In a parallel with a more traditional regression framework, this is equivalent to controlling for province, urban area and sector fixed effects, so to compare the treated unit only within the group of firms that share the same economic environment and are subject to the same shocks.[17] Moreover, the synthetic control routine does not allow for gaps in the time series of the relevant variables, so that our sample only includes treated and donor firms that are present in all years between 2005 and 2014.

Based on this pool of potential donors, we compute the aggregation weights for our synthetic controls; they are obtained as the values that minimize a penalty function over the pre-treatment values of the outcome variable (which in our case is again value added per worker) and other relevant covariates. In our case, the covariates include log sales, employees and the firm's age; the first two are averaged over the 5 years before the treatment, while the latter is considered at time $t-1$. On the outcome's variable side, we consider the values of the log value added per worker at $t-5$ and $t-1$. Using this procedure, synthetic firms are obtained as linear combinations of selected donor firms; they are the most similar object —both in terms of observables and non-observables— to the treated units of interest in the pre-treatment period, and this allows us to claim that they represent a reasonable counterfactual for the evolution of the outcome variable in the absence of treatment.

After this step, we end up with a synthetic control firm for every treated unit in our sample, together with the Root Mean Square Prediction Error (RMSPE) associated with the pre-treatment trajectory of the outcome variable in comparison with that of the treated firm. More precisely, we found a synthetic control for 38,345 out of 43,831 firms relocating between 2010 and 2012 (the 87.5%); for the remaining ones, either the optimization procedure wasn't able to achieve a convergence or the potential pool of donors was empty.

Table 3.12 contains the average values of some observables for treated and control units, in order to check for the pre-treatment balancing properties of our two groups of firms. Synthetic controls represent a very close match for treated firms in terms of the observables. The recorded discrepancies are most of the times very small, apart from a marginally-significant difference in the wage bill and a more sizable one in terms of age: on average, the treated firms are 1.3 years younger than the synthetic controls in the year before the switch takes place. This difference means that the procedure wasn't able to find —within the pool of donors— sufficiently young firms that

[17] Restricting the pool of potential donors has also the beneficial effect of reducing the computation time, which is a non-negligible issue, given the computationally-intensive nature of the optimization procedure that searches for the weights. When, after applying the described conditions, the pool of donors is still too large, we restrict it further by progressively imposing the same legal form, the same LLM and the same 4-digits sector as the treated unit.

matched the good performance of relocating firms in terms of productivity, sales and employees: in other words, all of these characteristics have been satisfactorily approximated by the synthetic control method, at the cost of including slightly more mature firms among the donors. We will be returning on this issue of age in the discussion of the results.

Finally, we are left with the task of re-aggregating the evolution of all of our treated-control pairs, in order to quantify the average effect of relocation. The empirical literature on synthetic controls usually focuses on a single case study, and only few of them deal with a set of multiple treated units (Dube and Zipperer, 2015; Kreif et al., 2016). In what follows, we build on the method outlined by Mirenda et al. (2017) to aggregate our results. We first compute the average log productivity for the two groups of switchers and synthetic controls, distinguishing by the year in which the move takes place. We weight the observations by the inverse of the Root Mean Squared Prediction Error (RMSPE), so that we give a greater importance to the synthetic controls that more closely follow their corresponding treated units in the pre-treatment period; to exclude outliers, we drop the units lying in the top or bottom 1% of the RMSPE distribution. Figure 3.6 graphically displays the results, depicting the evolution of average log productivity for treated and synthetic controls for each switch year; the horizontal axis represents the distance in years from the moment of the switch (we conventionally set the relocation year to zero). First of all, it is important to appreciate that the pre-treatment evolution of log productivity is remarkably similar between switching firms and synthetic controls, irrespective of the year in which relocation takes place. The effect of relocation, instead, shows different patterns, depending on the year of the switch: firms moving in 2010 and 2011 start accruing a productivity gain in the first year after the switch, and such gain is increasing with time; in the case of firms switching in a year of severe downturn such as 2012, instead, the decision to switch seems to entail a productivity loss in the first years, and turns into a gain only at a two-years distance from the move.

This evidence points to the fact that the size and timing of the productivity gain may depend on general economic conditions at the time of the switch and in the following years. As stated above, the fact of observing units being treated in different years is useful to net from year-specific effects that may affect the size of the gain. To do that, we pool together all the observations, irrespective of the switch year, and separately estimate the following simple model for treated firms and synthetic controls:

$$\pi_{i(s)y} = \delta_y + \sum_{t=-7}^{4} \gamma_{st} I_{iy}^{t} + \varepsilon_{i(s)y}, \text{ for } s \in \{T, C\} \tag{3.4}$$

where $\pi_{i(T)y}$ is the log productivity of treated firm i in year y, while $\pi_{i(C)y}$ is the log productivity of the synthetic control for firm i in year y; as above, δ_y are year-fixed effects. I_{iy}^t is equal to one if firm i in year y is t years away from the switch; negative values of t refer to the years before the relocation, that happens at $t = 0$. The summation runs from -7 to 4 which is the range of time periods spanned by our synthetic control exercise. Observations are once more weighted by the inverse of the RMSPE, in order to give more relevance to observations with a greater information content.

We are especially interested in the coefficients γ_{Tt} and γ_{Ct}, that capture —respectively for treated and controls— the average log labor productivity at a t-years distance from the relocation episode, net of year-specific effects common to all firms. For our synthetic control exercise to be meaningful, we would expect the coefficients for treated and synthetic controls to stay very close until the time of the switch; any subsequent divarication between the two series can be interpreted as the effect of the relocation on productivity. Figure 3.7 displays the estimates of the coefficients γ_{st}. As expected, the two series remain very similar until the time of the relocation; the productivity of the two groups starts to diverge form the first year after the switch, and the distance keeps increasing in the following years. At a 4-years distance from the move, the productivity gain for switchers is about 10 percentage points.

To draw a parallel with the difference-in-differences exercise presented in the previous section, we now want to break up this average effect, exploring the hypothesis that the size of the gain may vary according to the urban nature of the LLMs of departure and arrival. To do that, we estimate the following model, using as a dependent variable the log productivity difference between treated and controls:[18]

$$\Delta\pi_{iy} \equiv \pi_{i(T)y} - \pi_{i(C)y} = \delta_y + \sum_{t=-7}^{4} \gamma_t^G \left(\delta_G \times I_{iy}^t\right) + \varepsilon_{iy} \tag{3.5}$$

where δ_G is a set of fixed effects that capture the nature of the transition (e.g. from non-urban to urban or viceversa), like in the D-i-D exercise. Under this formulation, the coefficients γ_t^G directly measure the productivity gain (or loss) that firms in the transition group G experience at a distance of t years from the switch. For our exercise to be meaningful, we would expect two things to happen. First of all, the gain should be close to zero in the absence of treatment (that is, for negative values of t), meaning that the productivity of treated and controls evolves in a similar fashion. In the second place, as argued above, we would expect the productivity gain to be positive for

[18]This is equivalent to separately estimating the model on the log productivity of treated firms and synthetic controls, and then taking the difference of the estimated γ_t^G coefficients.

all kind of transitions, albeit with different intensities: being the result of the decision of supposedly rational agents, any type relocation should —on average— make the switching firm better off.

Figure 3.8 shows that these two expectations are met in our exercise.[19] As in the difference-in differences exercise, firms moving from a non-urban to a urban environment get the greatest productivity gain from relocating; this is only partly acquired upon switching (at $t = 0$), while the most of it is accrued in the following years, on a steadily increasing pattern. On the contrary, firms that downgrade from a urban to a non-urban area suffer on impact a productivity loss (switching costs more than compensate the benefits), despite recovering and turning positive in the following years; their productivity profile remains, however, constantly below the one of the other firms. In between these two extremes lie the firms that —albeit relocating— remain in the same type of LLM (either urban or non-urban). For these firms, the productivity gain is absent at the time of relocation, rapidly grows in the following two years and then remains roughly stable.

Further insights can be gained from the inspection of Figure 3.9, where we add the distinction between small and big urban areas. Firms that "upgrade" to a big urban area (either from a small or a non-urban one) benefit on impact from a productivity gain of about 5 percentage points; specularly, the firms that move away from a big urban area to a non-urban one suffer a loss. This hints at the existence of a static productivity premium that is specific to big urban areas only: the simple fact of being located in a big urban LLM entails an advantage that is lost once the firm moves away. On the dynamic side, productivity advantages also accrue more rapidly to upgrading firms (including those moving from non-urban to small urban LLMs), although with some swings. The firms that acquire the lowest gain are those that "downgrade" to smaller LLMs, while the switchers that remain in the same type of LLM again lie in between these two extremes.

Moreover, we address the issue of the imperfect balancing of characteristics discussed in Table 3.12. To do that, we add controls for age and size to the specification in equation 3.5. That should not only help us in rebalancing the observable characteristics across the two groups, but also in controlling for potential unobservables that are correlated with these two variables. Results are graphically displayed in Figures 3.10 and 3.11. The broad picture does not change and relative ranks are preserved; only the size of the gain slightly reduce over the 4-years horizon.

Finally, in Table 3.13 we provide a quantitative comparison of the size of productivity gains obtained with the difference-in-differences and with the

[19]The black dashed line represents the difference between the two lines in Figure 3.7.

synthetic controls methods. The comparison is offered at two different points in time, at the relocation year and 4 years after. The bulk of the differences between the two approaches lies in the estimate of the gain in the year of the switch. D-i-D generally estimates positive relocation premia, with a wide heterogeneity depending on the type of relocation that takes place; on the contrary, synthetic controls show much smaller —and even negative— effects on impact. The biggest gains are reaped by firms upgrading to a urban or a bigger LLM, that also accumulate more rapidly a productivity advantage in the following years. After four years, the gain accrued to this group of firms amounts to more than 15 percentage points, irrespective of the estimation method. The synthetic control exercise returns significantly different results from the D-i-D strategy for the group of firms that downgrade to a smaller or non-urban LLM: after four years their productivity gain does not exceed the 5 percentage points.

Overall, these results point to the existence of a static productivity premium, in particular for bigger cities. The steeper productivity path displayed by the firms moving to a urban environment, though, suggests that learning channels are also at work in cities, irrespective of their size. Further evidence on the existence of these broad patterns is provided by the empirical exercise presented in the appendix, where we borrow from the regression framework that De La Roca and Puga (2017) use to estimate the effect of urban experience on wages.

3.6 Concluding remarks

In this paper we study the determinants of urban productivity premium, disentangling the relative contribution of two potential explanations that the literature lists among the most relevant factors in shaping the cities' advantage: on the one hand, the sorting of more productive firm to bigger cities; on the other hand, the positive externalities arising from agglomeration economies. To do that, we exploit a unique database containing the universe of Italian firms in the period 2005–2014. The key feature of this data is that we are able to observe the location of the firms at each point in time, and therefore we are able to identify firms relocating from one city to another.

In an empirical framework similar to the one of Abowd et al. (1999), this feature allows us to separately identify two components of firm productivity: the first one captures the intrinsic quality of the firms, while the second one is specific to each city and gauges the extent of the agglomeration externalities. Doing so, we are able to measure the relative importance of sorting and agglomeration in determining the urban productivity premium we observe in the aggregate.

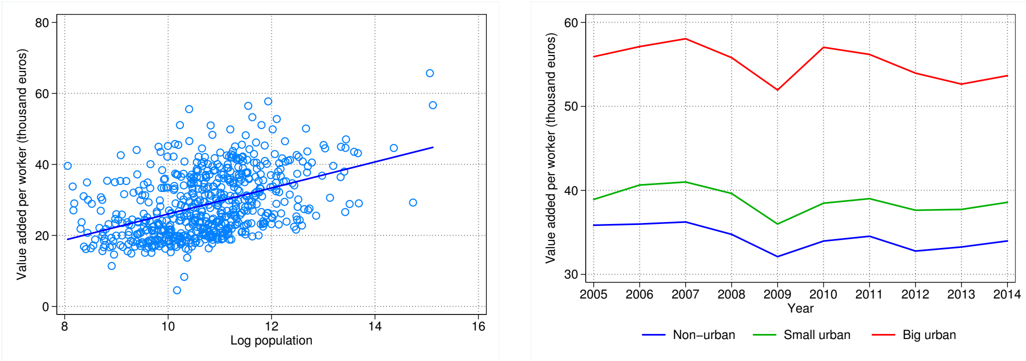
Our results show that —conditional on observable characteristics of the firms— urban areas display on average a productivity premium of roughly 15 percentage points with respect to non-urban locations. The intrinsic qualities of the firms —as captured by the firm-specific component— explains a substantial fraction of the variability in firm-level productivity; the variability commanded by the city-specific component, instead, is much lower, despite being non-negligible. Both firm- and city-specific effects display a positive elasticity with respect to city size, with the former being higher than the latter. This means on the one hand that firms with higher intrinsic qualities sort into bigger cities (firms that sort into urban areas are on average 7 percentage points more productive than the others); on the other hand, it means that agglomeration economies are stronger for urban areas.

We then explore the mechanisms through which firms appropriate the advantages of agglomeration economies in urban areas. In order to do so, we exploit once more the fact that we observe firms relocating across space; we set up two empirical exercised based on difference-in-differences and synthetic controls, with the aim of observing the differential pattern of productivity accrual across different destinations. Relocation is associated with significant productivity gains, on average equal to 10 percentage points at a 4-years distance from the switch; the gain is highest for firms relocating from non-urban areas to urban ones, while it is lowest for firms going the other way round. Part of the gain accrues immediately after the switch, but only for firms moving to big urban areas; the nature of this static advantage is however transitory, since firms moving away from big cities suffer a specular productivity drop. The most relevant channel of productivity accrual is, however, the dynamic one. Relocation gains accumulate over time, though at higher rates for firms that move from non-urban to urban locations. This is consistent with faster learning processes, that entail a higher reward for an extra year of experience spent in a urban environment.

Figures and Tables

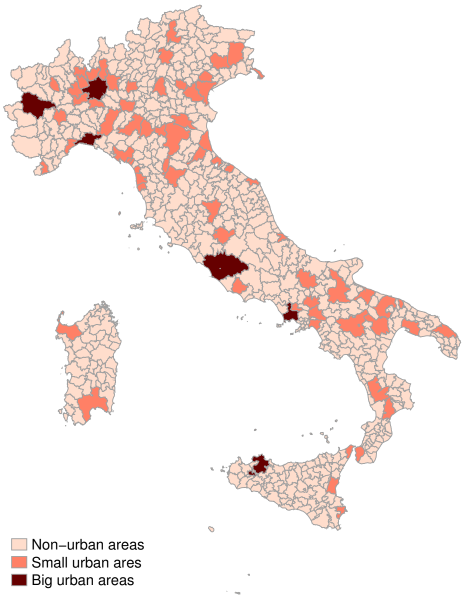
Figure 3.1: The urban productivity premium in Italy

a) Productivity vs. city size *b) Productivity across urban areas*



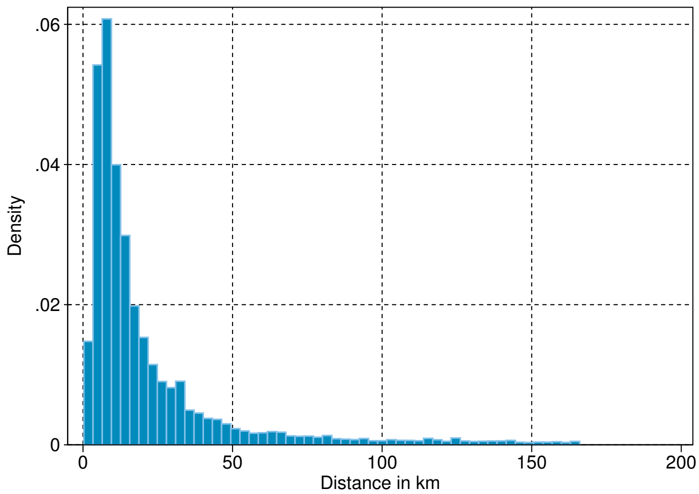
Notes: In panel (a) each dot represents a Local Labour Market (LLM), as defined by Istat. Large urban areas are defined as those with more than 500 thousand inhabitants.

Figure 3.2: The geography of Italian urban areas



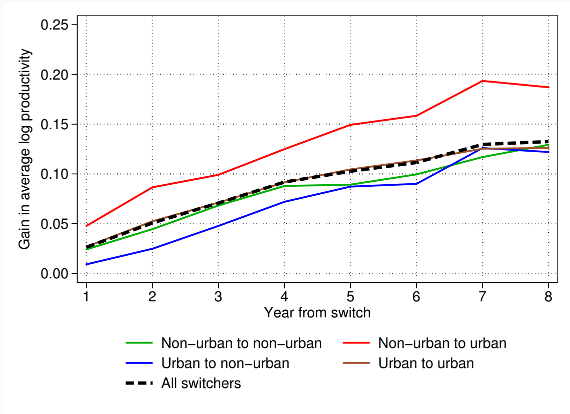
Notes: Own elaborations on Istat and Eurostat data.

Figure 3.3: Histogram of relocation distances



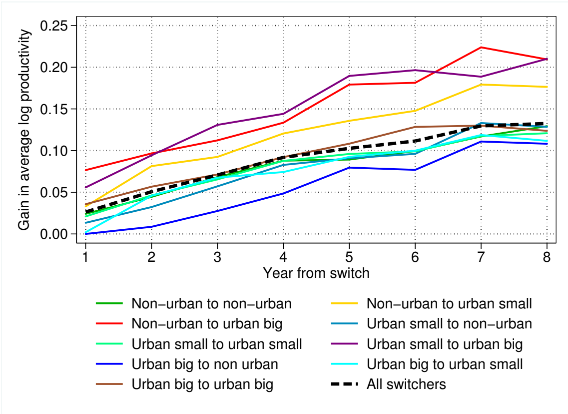
Notes: Data have been truncated at the 90th percentile of the observed distribution of distances.

Figure 3.4: Difference-in-difference – Productivity gain upon relocation



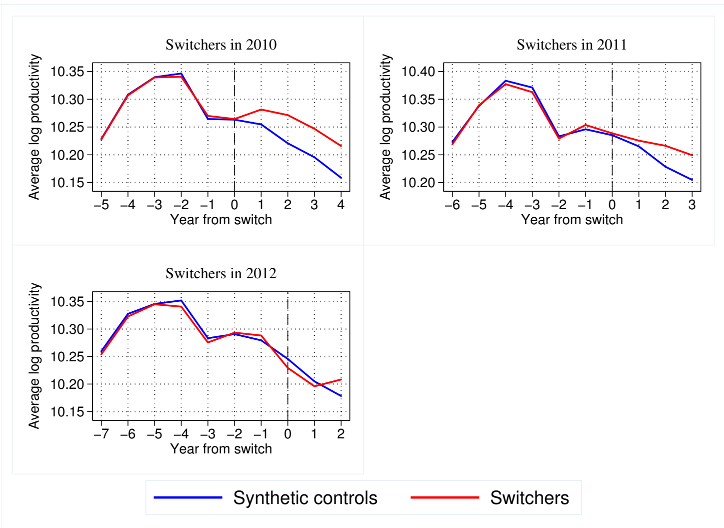
Notes: The figure displays the coefficient estimates from the difference-in-difference specification, by type of relocation and by the number of years after the move. The year in which the relocation takes place is set to zero. Quantitative estimates and significances are displayed in Table 3.15.

Figure 3.5: Difference-in-difference – Productivity gain upon relocation



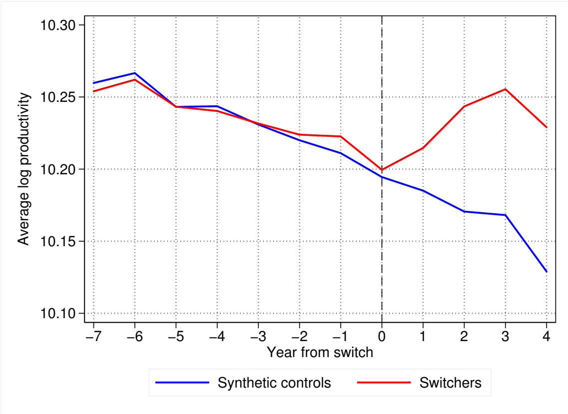
Notes: The figure displays the coefficient estimates from the difference-in-difference specification, by type of relocation and by the number of years after the move. The year in which the relocation takes place is set to zero. Quantitative estimates and significances are displayed in Table 3.16.

Figure 3.6: Synthetic controls, results by year of switch



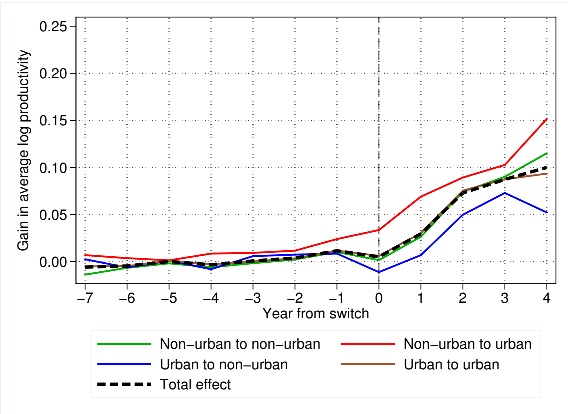
Notes: The figure displays the average log productivity of treated firms (i.e. movers) against that of the synthetic controls, by year of switch. The year in which the relocation takes place is set to zero. Observations are weighted by the inverse of the Root Mean Square Prediction Error (RMSPE). To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Figure 3.7: Synthetic controls, pooled years



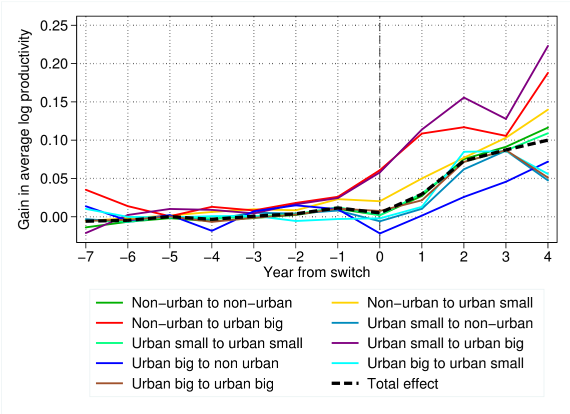
Notes: The figure displays the coefficient estimates from the synthetic control exercise. The year in which the relocation takes place is set to zero. To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Figure 3.8: Synthetic controls, by type of switch



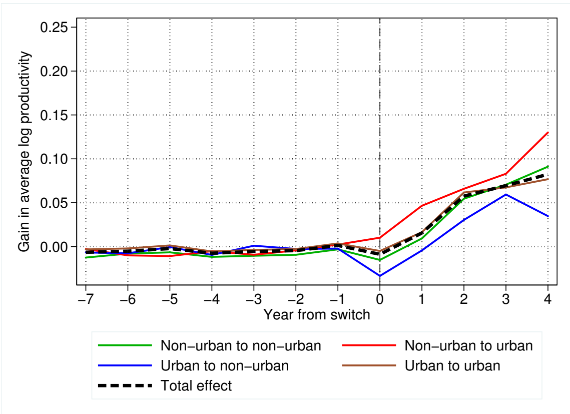
Notes: The figure displays the coefficient estimates from the synthetic control exercise, by type of relocation and by the number of years before and after the move. The year in which the relocation takes place is set to zero. To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Figure 3.9: Synthetic controls, by type of switch (detail)



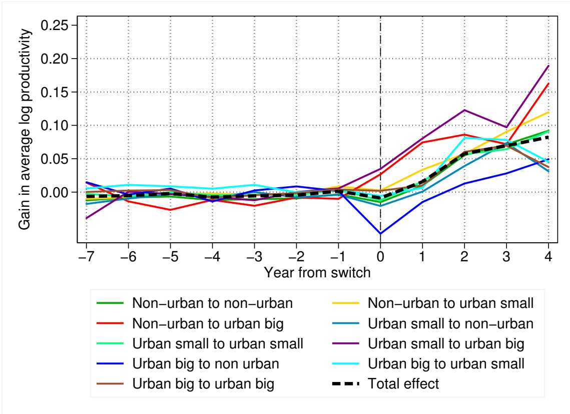
Notes: The figure displays the coefficient estimates from the synthetic control exercise, by type of relocation and by the number of years before and after the move. The year in which the relocation takes place is set to zero. To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Figure 3.10: Synthetic controls netting for age and size, by type of switch



Notes: The figure displays the coefficient estimates from the synthetic control exercise, by type of relocation and by the number of years before and after the move. The year in which the relocation takes place is set to zero. To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Figure 3.11: Synthetic controls netting for age and size, by type of switch (detail)



Notes: The figure displays the coefficient estimates from the synthetic control exercise, by type of relocation and by the number of years before and after the move. The year in which the relocation takes place is set to zero. To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution.

Table 3.1: Population growth in urban areas in selected countries

	Growth rate of population (10-years averages)		
	time range	urban areas	national average
USA	1920–2010	17.9	12.6
Spain	1920–2010	18.1	8.9
France	1937–2007	7.7	5.7
Italy	1911–2001	7.2	4.9

Source: Giffoni et al. (2016), Duranton and Puga (2014).

Table 3.2: Evolution of Italian LLMs over time

Census year	Urban	Non Urban	Totale	LLM definition
1981	74	880	954	1981
1991	74	710	784	1991
2001	74	612	686	2001
2001	74	609	683	2011
2011	74	538	612	2011

Source: Istat, Census data in 1981, 1991, 2001 and 2011.

Table 3.3: Descriptive statistics

	Non-urban			Urban		
Year	# firms	VA per worker	# workers per firm	# firms	VA per worker	# workers per firm
2005	1,851,819	35,850	3.33	2,442,237	46,700	4.10
2006	1,866,562	35,987	3.36	2,463,383	48,199	4.15
2007	1,902,740	36,234	3.41	2,496,657	48,770	4.18
2008	1,914,794	34,761	3.44	2,513,636	47,010	4.23
2009	1,890,381	32,108	3.40	2,488,411	43,295	4.20
2010	1,882,643	33,964	3.36	2,484,337	46,971	4.17
2011	1,871,957	34,535	3.37	2,483,459	46,919	4.17
2012	1,869,098	32,774	3.34	2,478,877	45,159	4.16
2013	1,840,623	33,256	3.31	2,455,844	44,665	4.15
2014	1,819,558	33,978	3.27	2,444,285	45,656	4.13

Source: own elaborations from Archivio statistico delle imprese attive (ASIA).

Table 3.4: The urban productivity premium

	(1)	(2)	(3)	(4)	(5)	(6)
log LLM population	0.0558***	0.0372***				
	(0.0002)	(0.0002)				
age		0.0114***		0.0112***		0.0113***
		(0.0017)		(0.0017)		(0.0017)
age^2		-0.0001***		-0.0001***		-0.0001***
		(0.0000)		(0.0000)		(0.0000)
Urban area			0.1530***	0.0933***		
			(0.0007)	(0.0006)		
Small urban area					0.1360***	0.0760***
					(0.0008)	(0.0007)
Big urban area					0.1770***	0.1210***
					(0.0009)	(0.0009)
Year FE	Y	Y	Y	Y	Y	Y
Sector FE	Y	N	Y	N	Y	N
Size class FE	Y	N	Y	N	Y	N
Obs.	41,295,233	41,295,232	41,295,233	41,295,232	41,295,233	41,295,232
R^2	0.012	0.218	0.010	0.216	0.010	0.217

Notes: The dependent variable is log value added per worker at the firm level. Standard errors clustered at the firm level in parentheses. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.5: Correlation table for firm and city fixed effects

	π	δ_c	$\delta_{c,TV}$	δ_f
δ_c	0.1460			
$\delta_{c,TV}$	0.1591	0.8983		
δ_f	0.7305	0.0456	0.0454	
$\delta_{f,TV}$	0.7606	0.0464	0.0389	0.9935

Notes: Variable π is log value added per worker. The fixed effects labeled with the TV subscript are those obtained from the model with time-varying city fixed effects.

Table 3.6: City-specific components and the size of urban areas

	(1)	(2)	(3)
log LLM population	0.0191***		
	(0.00527)		
Urban area		0.0415***	
		(0.0150)	
Small urban area			0.0317***
			(0.0149)
Big urban area			0.0506***
			(0.0217)
Obs.	79,404	79,404	79,404
R^2	0.042	0.037	0.037

Notes: The dependent variable is the city-specific fixed effect estimated according to equation 3.1. Year fixed effects included in all specifications. Standard errors clustered at the LLM level in parentheses. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.7: Firm-specific components and the size of urban areas

	(1)	(2)	(3)
log LLM population	0.0239***		
	(0.000221)		
Urban area		0.0739***	
		(0.000679)	
Small urban area			0.0721***
			(0.000771)
Big urban area			0.0766***
			(0.000884)
Obs.	6,296,667	6,296,667	6,296,667
R^2	0.002	0.002	0.002

Notes: The dependent variable is the firm-specific fixed effect estimated according to equation 3.1. Relocating firms have been dropped from the estimates in order to have a univocal association with a LLM across time. Standard errors clustered at the LLM level in parentheses. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.8: Transition matrix for relocating firms

from \ to	Non-urban	Urban small	Urban big
Frequencies (# of relocation episodes)			
Non-urban	154,628	44,018	24,986
Urban small	45,003	149,663	21,914
Urban big	23,859	20,277	108,662
Transition matrix (empirical probabilities)			
Non-urban	0.6914	0.1968	0.1117
Urban small	0.2078	0.6910	0.1012
Urban big	0.1561	0.1327	0.7111

Notes: Authors' computation on data from Archivio statistico delle imprese attive (ASIA). Big urban areas are defined as those with more than 500 thousand inhabitants; small urban areas are the remaining ones.

Table 3.9: Descriptive statistics for relocating firms

	Averages		
	Movers	Stayers	Difference
Employees	6.03	3.64	2.39***
Sales	1,489	629	860***
Age	9.37	13.60	-4.22***
VA per worker	35.67	26.76	8.91***

Notes: Authors' computation on data from Archivio statistico delle imprese attive (ASIA). Sales and VA per worker are expressed in thousand Euros at constant 2010 prices. For the moving firms, only the period before the switch has been considered in the computation of the average values. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.10: The productivity premium of relocating firms

	(1)	(2)	(3)	(4)
Switcher	0.0818***			
	(0.00155)			
t=-9 years		0.0852***		
		(0.00603)		
t=-8 years		0.0817***		
		(0.00405)		
t=-7 years		0.0863***		
		(0.00331)		
t=-6 years		0.0897***		
		(0.00248)		
t=-5 years		0.0870***		
		(0.00229)		
t=-4 years		0.0869***		
		(0.00211)		
t=-3 years		0.0799***		
		(0.00201)		
t=-2 years		0.0695***		
		(0.00188)		
t=-1 year		0.0779***		
		(0.00177)		

continues on next page

The productivity premium of relocating firms (continued)

	(1)	(2)	(3)	(4)
Non-urban to non-urban			0.0348***	0.0347***
			(0.00222)	(0.00222)
Non-urban to urban			0.0683***	
			(0.00361)	
Urban to non-urban			0.0546***	
			(0.00359)	
Urban to urban			0.113***	
			(0.00195)	
Non-urban to urban small				0.0588***
				(0.00439)
Non-urban to urban big				0.0848***
				(0.00576)
Urban small to non-urban				0.0468***
				(0.00428)
Urban small to urban small				0.0928***
				(0.00233)
Urban small to urban big				0.171***
				(0.00713)
Urban big to non-urban				0.0696***
				(0.00610)
Urban big to urban small				0.139***
				(0.00704)
Urban big to urban big				0.128***
				(0.00304)
Obs.	39,365,734	39,187,638	39,187,638	39,187,638
R^2	0.232	0.231	0.231	0.231

Notes: Authors' computation on data from Archivio statistico delle imprese attive (ASIA). The estimates displayed in columns (2)-(4) are obtained excluding firms relocating multiple times. For relocating firms, we only consider the years before the switch takes place. In all regressions, controls include: age, age squared, sector (4 digit) FE, year FE, size class FE and legal form FE. Standard errors clustered at the firm level in parentheses. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.11: Difference-in-difference regressions on switching premium

	(1)	(2)	(3)
Switch	0.0675***		
	(0.00155)		
Non-urban to non-urban		0.0620***	
		(0.00277)	
Non-urban to urban		0.100***	
		(0.00524)	
Urban to non-urban		0.0480***	
		(0.00502)	
Urban to urban		0.0681***	
		(0.00215)	
Non-urban to non-urban			0.0620***
			(0.00277)
Non-urban to urban small			0.0919***
			(0.00627)
Non-urban to urban big			0.117***
			(0.00949)
Urban small to non-urban			0.0553***
			(0.00603)
Urban small to urban small			0.0620***
			(0.00279)
Urban small to urban big			0.116***
			(0.0104)
Urban big to non urban			0.0324***
			(0.00900)
Urban big to urban small			0.0549***
			(0.00962)
Urban big to urban big			0.0724***
			(0.00375)
Firm FE	Y	Y	Y
Year FE	Y	Y	Y
Obs.	20,145,562	20,145,562	20,145,562
R^2	0.725	0.725	0.725

Notes: The dependent variable log value added per worker at the firm level. Standard errors in brackets clustered at the firm level. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.12: Balancing properties for the synthetic control strategy

	Averages		
	Movers	Synthetic controls	Difference
log VA per worker ($t-1$)	10.292	10.285	0.0074
log VA per worker ($t-5$)	10.312	10.312	-0.0004
Employees ($t-1$)	7.95	7.21	-0.7358
Wage bill ($t-1$)	26,852	25,467	1,386*
Age ($t-1$)	15.79	17.05	-1.26***
log sales ($t-1$)	12.21	12.12	0.0898
Obs.	37,578	37,578	

Notes: Averages are computed on the pre-treatment period reported in parentheses (t is the year of relocation). Wage bill is expressed in thousand Euros. Observations are weighted by the inverse of the Root Mean Square Prediction Error (RMSPE). To exclude outliers, data have been trimmed from the top and bottom 1% of the RMSPE distribution. Significance level: * p<0.10, ** p<0.05, *** p<0.01.

Table 3.13: Comparing relocation gains across identification strategies

	Diff-in-diff		Synthetic controls		Synthetic controls (net of age and size)	
	Switch year	After 4 years	Switch year	After 4 years	Switch year	After 4 years
Non-urban to non-urban	0.0241	0.0892	0.0016	0.1151	-0.0152	0.0911
Non-urban to urban	0.0477	0.1494	0.0336	0.1516	0.0102	0.1301
Urban to non-urban	0.0091	0.0873	-0.0111	0.0522	-0.0334	0.0347
Urban to urban	0.0268	0.1045	0.0062	0.0935	-0.0050	0.0768
Non-urban to non-urban	0.0241	0.0892	0.0020	0.1164	-0.0150	0.0920
Non-urban to urban small	0.0331	0.1357	0.0203	0.1399	0.0024	0.1197
Non-urban to urban big	0.0766	0.1792	0.0609	0.1878	0.0267	0.1626
Urban small to non-urban	0.0134	0.0908	-0.0059	0.0475	-0.0207	0.0312
Urban small to urban small	0.0210	0.0960	0.0019	0.1091	-0.0122	0.0910
Urban small to urban big	0.0558	0.1896	0.0579	0.2230	0.0348	0.1893
Urban big to non-urban	0.0000	0.0796	-0.0218	0.0720	-0.0622	0.0492
Urban big to urban small	0.0021	0.0925	-0.0021	0.0559	-0.0057	0.0454
Urban big to urban big	0.0357	0.1083	0.0076	0.0511	0.0019	0.0375

Appendix

The effect of urban experience

As an additional evidence on the mechanism by means of which firms accrue different productivity advantages in different urban environments, we build on the empirical approach of De La Roca and Puga (2017) to uncover the effect of urban experience on the observed productivity premium. To do that, we are forced to restrict our sample to the subset of firms for which we are able to quantify —at every point in time— the number of years spent in each location. As a consequence, we focus on the group of nearly 4 million firms born from 2005 on. Despite being a highly selected group of young firms, the estimated urban productivity premium is roughly the same as the one obtained with the full sample, conditional on observables.

In our empirical exercise we disentangle the "static" productivity premium arising from being located in a urban area from the "dynamic" effect related to the amount of experience accumulated in those urban environments. We estimate the following model:

$$\begin{aligned}\pi_{iy} =&\alpha_1 US_{iy} + \alpha_2 UB_{iy} + \beta_0 age_{iy} + \beta_1 ExpUS_{iy} + \beta_2 ExpUB_{iy} + \\ &+\gamma_0 age_{iy}^2 + \gamma_1 ExpUS_{iy} \times age_{iy} + \gamma_2 ExpUB_{iy} \times age_{iy} + \\ &+\delta_y + \delta_s + \delta_e + \delta_i + \varepsilon_{iy}\end{aligned} \tag{3.6}$$

where π_{iy} is again the log value added per worker of firm i in year y. The time-varying dummies US and UB indicate whether the firm is located in a small urban area or in a big urban area, respectively. Similarly, variables $ExpUS$ and $ExpUB$ capture the number of years spent by the firm in a small or in a big urban area. The δ variables stand for year, sector, size class and firm fixed effects. The latter is only included in certain specifications to partial out the sorting behaviour of firms, in the spirit of the exercise performed in section 3.4.[20] Since the age is by definition the sum of the experiences accumulated in all environments (non-urban, small and big urban areas), the excluded category is always represented by the firms located in non-urban areas: the α coefficients capture the productivity advantage of being located in urban areas, with respect to non-urban areas; similarly, while β_0 captures the effect of one more year of experience in a non-urban area, the other β coefficients gauge the additional return from accruing that year in a urban environment; the same argument runs for the interaction terms, that

[20]When controlling for firm fixed effects, year fixed effects have to be dropped to avoid collinearity with the age term.

capture the (differential) curvature of the productivity reward to experience in every urban environment.

Results are presented in Table 3.14 for different specifications, in which we progressively add firm-level controls. All variables are strongly significant across specifications. The urban productivity premium emerges clearly and is increasing in the size of the urban area; it suffers a sharp drop when we net out the sorting behaviour of firms by including firm fixed effects, consistently with the evidence presented in section 3.4. The productivity premium also has a dynamic component that accrues according to the number of years spent in a urban environment, as shown by the positive coefficients attached to the experience terms; the returns to an extra year of experience is higher in a urban environment, with a twofold premium for the experience acquired acquired in a big urban area with respect to a small urban one. As shown by the coefficients on the interaction terms, productivity is characterized by decreasing returns to experience; the curvature is increasing in the size of the urban area where the experience is accrued.

Additional tables

Table 3.14: Productivity premium of urban firms

	(1)	(2)	(3)
Urban Small	0.0831***	0.0453***	0.00855**
	(0.00192)	(0.00169)	(0.00344)
Urban Big	0.122***	0.0729***	0.0327***
	(0.00217)	(0.00193)	(0.00432)
age	0.0474***	0.0376***	0.000887**
	(0.000444)	(0.000402)	(0.000391)
age^2	-0.00137***	-0.00103***	-0.00121***
	(0.0000325)	(0.0000292)	(0.0000253)
Experience Urban Big	0.0128***	0.0112***	0.0133***
	(0.000795)	(0.000709)	(0.000684)
age $\times$ Experience Urban Big	-0.000405***	-0.000390***	-0.000653***
	(0.0000584)	(0.0000520)	(0.0000451)
Experience Urban Small	0.00935***	0.00600***	0.00603***
	(0.000684)	(0.000608)	(0.000589)
age $\times$ Experience Urban Small	-0.000337***	-0.000227***	-0.000279***
	(0.0000494)	(0.0000440)	(0.0000384)
Industry FE	Y	Y	Y
Size class FE	Y	Y	Y
Year FE	Y	Y	N
Firm FE	N	N	Y
Obs.	16,544,078	16,544,076	15,887,432
R^2	0.021	0.194	0.657

Notes: Authors' computation on data from Archivio statistico delle imprese attive (ASIA). Clustered SE in parenthesis at firm level. * significant at 10 per cent, ** at 5 per cent, *** at 1 per cent.

Table 3.15: Difference-in-difference regressions, by number of years after switch

Years after switch	All switchers	Non-urban to-non urban	Non-urban to urban	Urban to non-urban	Urban to urban
0	0.0263	0.0241	0.0477	0.0091*	0.0268
	(0.0017)	(0.0031)	(0.0059)	(0.0056)	(0.0024)
1	0.0507	0.0445	0.0865	0.0248	0.0524
	(0.0019)	(0.0035)	(0.0065)	(0.0063)	(0.0027)
2	0.0707	0.0683	0.0991	0.0477	0.0712
	(0.0021)	(0.0038)	(0.0071)	(0.0068)	(0.0029)
3	0.0918	0.0879	0.1249	0.0720	0.0916
	(0.0022)	(0.0039)	(0.0073)	(0.0072)	(0.0031)
4	0.1027	0.0892	0.1494	0.0873	0.1045
	(0.0025)	(0.0044)	(0.0084)	(0.0081)	(0.0034)
5	0.1115	0.0995	0.1585	0.0900	0.1136
	(0.0027)	(0.0049)	(0.0093)	(0.0090)	(0.0038)
6	0.1296	0.1169	0.1935	0.1260	0.1254
	(0.0031)	(0.0055)	(0.0102)	(0.0101)	(0.0043)
7	0.1325	0.1293	0.1871	0.1221	0.1261
	(0.0036)	(0.0062)	(0.0125)	(0.0120)	(0.0049)

Notes: The dependent variable log value added per worker at the firm level. The year in which the relocation takes place is set to zero. Standard errors in brackets clustered at the firm level. The asterisk marks the coefficients that are not statistically significant at least at the 95% level.

Table 3.16: Difference-in-difference regressions, by number of years after switch (detail)

Years after switch	Non-urban to non-urban	Non-urban to urban small	Non-urban to urban big	Urban small to non-urban	Urban small to urban small	Urban small to urban big	Urban big to non-urban	Urban big to urban small	Urban big to urban big
0	0.0241	0.0331	0.0766	0.0134	0.0210	0.0558	0.0000*	0.0021*	0.0357
	(0.0031)	(0.0071)	(0.0106)	(0.0067)	(0.0031)	(0.0117)	(0.0102)	(0.0113)	(0.0041)
1	0.0445	0.0814	0.0966	0.0323	0.0463	0.0943	0.0086*	0.0458	0.0567
	(0.0035)	(0.0077)	(0.0122)	(0.0075)	(0.0034)	(0.0132)	(0.0113)	(0.0118)	(0.0047)
2	0.0683	0.0924	0.1123	0.0571	0.0655	0.1308	0.0275	0.0676	0.0714
	(0.0038)	(0.0084)	(0.0130)	(0.0081)	(0.0037)	(0.0141)	(0.0123)	(0.0127)	(0.0051)
3	0.0879	0.1205	0.1335	0.0828	0.0877	0.1441	0.0486	0.0743	0.0922
	(0.0039)	(0.0087)	(0.0136)	(0.0087)	(0.0040)	(0.0150)	(0.0127)	(0.0142)	(0.0055)
4	0.0892	0.1357	0.1792	0.0908	0.0960	0.1896	0.0796	0.0925	0.1083
	(0.0044)	(0.0099)	(0.0155)	(0.0098)	(0.0044)	(0.0169)	(0.0143)	(0.0149)	(0.0061)
5	0.0995	0.1477	0.1813	0.0961	0.0991	0.1965	0.0769	0.1000	0.1284
	(0.0049)	(0.0110)	(0.0172)	(0.0108)	(0.0048)	(0.0189)	(0.0161)	(0.0171)	(0.0065)
6	0.1169	0.1792	0.2239	0.1331	0.1175	0.1886	0.1109	0.1187	0.1299
	(0.0055)	(0.0122)	(0.0187)	(0.0123)	(0.0055)	(0.0219)	(0.0175)	(0.0187)	(0.0075)
7	0.1293	0.1765	0.2094	0.1285	0.1208	0.2101	0.1082	0.1117	0.1237
	(0.0062)	(0.0150)	(0.0227)	(0.0147)	(0.0063)	(0.0255)	(0.0208)	(0.0214)	(0.0085)

Notes: The dependent variable log value added per worker at the firm level. The year in which the relocation takes place is set to zero. Standard errors in brackets clustered at the firm level. The asterisk marks the coefficients that are not statistically significant at least at the 95% level.

References

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

Abadie, A. and J. Gardeazabal (2003). The Economic Costs of Conflict: a Case Study of the Basque Country. *The American Economic Review 93*(1), 113–132.

Abbate, C. C., M. G. Ladu, and A. Linarello (2017). An Integrated Dataset of Italian Firms: 2005-2014. *Bank of Italy Occasional Papers 384*, 3–26.

Abowd, J. M., R. H. Creecy, and F. Kramarz (2002). Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Center for Economic Studies, US Census Bureau.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Ahrend, R., E. Farchy, I. Kaplanis, and A. C. Lembcke (2014). What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries. OECD Regional Development Working Papers 2014/5, OECD Publishing.

Andersson, F., S. Burgess, and J. I. Lane (2007). Cities, Matching and the Productivity Gains of Agglomeration. *Journal of Urban Economics 61*(1), 112 – 128.

Andersson, M., J. Klaesson, and J. P. Larsson (2014). The Sources of the Urban Wage Premium by Worker Skills: Spatial Sorting or Agglomeration Economies? *Papers in Regional Science 93*(4), 727–747.

Baum-Snow, N. and R. Pavan (2011). Understanding the City Size Wage Gap. *The Review of Economic Studies 79*(1), 88–127.

Behrens, K. and F. Robert-Nicoud (2015). Agglomeration Theory with Heterogeneous Agents. In G. Duranton, V. J. Henderson, and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, Volume 5, Chapter 4, pp. 171–245. Elsevier.

Bergeaud, A. and S. Ray (2017). Frictions in the Corporate Real-estate Market, Firms' Relocation and Employment. Mimeo.

Brouwer, A. E., I. Mariotti, and J. N. Van Ommeren (2004). The Firm Relocation Decision: An Empirical Investigation. *The Annals of Regional Science 38*(2), 335–347.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics 63*(2), 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica 80*(6), 2543–2594.

Dauth, W., S. Findeisen, E. Moretti, and J. Suedekum (2016). Spatial Wage Disparities: Workers, Firms, and Assortative Matching. Mimeo. Avalilable at `https://sites.google.com/site/jenssuedekum/home/working-papers`.

D'Costa, S. and H. G. Overman (2014). The Urban Wage Growth Premium: Sorting or Learning? *Regional Science and Urban Economics 48*(C), 168–179.

De La Roca, J. and D. Puga (2017). Learning by Working in Big Cities. *Review of Economic Studies 84*(1), 106–142.

Di Giacinto, V., M. Gomellini, G. Micucci, and M. Pagnini (2014). Mapping Local Productivity Advantages in Italy: Industrial Districts, Cities or both? *Journal of Economic Geography 14*(2), 365–394.

Dijkstra, L. and H. Poelman (2012). Cities in Europe: the New OECD-EC Definition. EC Regional Focus.

Dube, A. and B. Zipperer (2015). Pooling Multiple Case Studies using Synthetic Controls: An Application to Minimum Wage Policies. IZA Discussion Paper No. 8944.

Duranton, G. and D. Puga (2004). Micro-foundations of Urban Agglomeration Economies. In V. J. Henderson and J. F. Thisse (Eds.), *Handbook* *of Regional and Urban Economics*, Volume 4, Chapter 48, pp. 2063–2117. North-Holland.

Duranton, G. and D. Puga (2014). The Growth of Cities. In *Handbook of Economic Growth*, Volume 2 of *Handbook of Economic Growth*, Chapter 5, pp. 781–853. Elsevier.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014). Spatial Sorting. *Journal of Political Economy 122*(3), 554–620.

Gaubert, C. (2017). Firm Sorting and Agglomeration. Mimeo.

Giffoni, F., M. Gomellini, and D. Pellegrino (2016). Human Capital and Urban Growth in Italy 1981-2001. *Bank of Italy Occasional Papers 1127*, 3–40.

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics 46*(2), 254–277.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer City. *Journal of Economic Geography 1*(1), 27–50.

Glaeser, E. L. and D. C. Maré (2001). Cities and Skills. *Journal of Labor Economics 19*(2), 316–42.

Henderson, J. V. (2003). Marshall's Scale Economies. *Journal of Urban economics 53*(1), 1–28.

Knoben, J., L. Oerlemans, and R. Rutten (2008). The Effects of Spatial Mobility on the Performance of Firms. *Economic Geography 84*(2), 157–183.

Kreif, N., R. Grieve, D. Hangartner, A. J. Turner, S. Nikolova, and M. Sutton (2016). Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units. *Health Economics 25*(12), 1514–1528.

Lamorgese, A. R. and A. Petrella (2016). An Anatomy of Italian Cities: Evidence from Firm-level Data. *Bank of Italy Occasional Papers 362*, 3–30.

Lamorgese, A. R. and A. Petrella (2017). Italian Cities: Definition, Characteristics and Growth. Forthcoming in Bank of Italy Occasional Papers.

Linarello, A. and A. Petrella (2017). Productivity and Reallocation: Evidence from the Universe of Italian Firms. *International Productivity Monitor 32*, 116.

Macis, M. and F. Schivardi (2016). Exports and Wages: Rent Sharing, Workforce Composition, or Returns to Skills? *Journal of Labor Economics 34*(4), 945–978.

Manyika, J., J. Remes, R. Dobbs, J. Orellana, and F. Schaer (2012). Urban America: US Cities in the Global Economy. Report, McKinsey Global Institute.

Marshall, A. (1890). *Principles of Economics.* Macmillan.

Mirenda, L., S. Mocetti, and L. Rizzica (2017). The Real Effects of Ndrangheta: Firm-level Evidence. Forthcoming in Bank of Italy Working Papers.

Moretti, E. (2004). Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions. *The American Economic Review 94*(3), 656–690.

UN (2013). World Population Prospects. Technical report, Population division of the UN Department of economic and social affairs.

Yankow, J. J. (2006). Why do Cities Pay more? An Empirical Examination of some Competing Theories of the Urban Wage Premium. *Journal of Urban Economics 60*(2), 139 – 161.