Mafia, corruption, and the cost of road infrastructures in Italy

An exploratory investigation

Thesis presented by
Giorgio Ferroni

Supervisors
Amma Panin, Dario Musolino

Reader
Romain Houssa

Academic year
2022

In order to obtain the Joint Degree
Master 120 en sciences économiques, orientation générale à finalité approfondie
(UCLouvain/UNamur)
And
Laurea specialistica in Discipline Economiche e Sociali (Bocconi)

Economics School of Louvain/UCLouvain • Place Montesquieu 3 • 1348 Louvain-la-Neuve
Département des Sciences économiques/UNamur • Rempart de la Vierge 8 • 5000 Namur
Università Commerciale Luigi Bocconi• Via Roberto Sarfatti, 25 • 20136 Milano - Italy

Mafia, corruption, and the cost of road infrastructures in Italy: an exploratory investigation

Giorgio Ferroni*

Abstract

This paper analyses the impact of corruption and organised crime on unit costs of road-infrastructure works in Italy. I build on and expand an original database by ANAS, the main national public agency in charge of the road network, about projects in 19 out of 20 Italy's regions and focus on estimating a function to identify the main costs' drivers. I collect data at both regional and provincial level for each observation and take a regression of unit-costs on mafia presence and corruption as well as on a number of other covariates to control for factors' prices and environmental specificities. Results from different specifications consistently show that the variables of interest have a statistically not significant impact on unit-costs' variation for public works.

*This paper is my master thesis work for which I would like to thank my supervisors Amma Panin and Dario Musolino, and 2 anonymous engineers from ANAS who allowed me to make use of their data.

Contents

1 Introduction

There is widespread accord on the positive role infrastructures play in a country's economic success and numerous studies have tried to quantify the magnitude of their impact (Donaldson, 2018; Banerjee, Duflo and Qian, 2020; Shiferaw, Söderbom, Siba and Alemu, 2012; Limao and Venables, 2001).

Since the massive investments required to modernize and maintain a national infrastructural stock, usually undertaken by governments and public authorities, are the thriving ground for corruption and anticompetitive phenomena the role played by corruption in economic growth too has been thoroughly analyzed and produced a wide literature (Mauro, 1995; Pellegrini and Gerlagh, 2004; Mo, 2001; Bardhan, 1997). Even though most recent findings have shown a negative impact on growth, overall results divide researchers between those who view corruption as "sand" in the wheels of development, rising transaction uncertainty and costs, or "grease", potentially making transactions smoother whereas bureaucracy would hinder exchange and production.

Assuming that infrastructures positively affect development, the aim of this paper is therefore functional to the improvement of this latter debate. In these pages I try to give a contribution to our knowledge about organised crime and corruption's role in costs formation in the road-construction sector, through the case of the Italian network. Thanks to a novel dataset about costs for a variety of works carried out by ANAS[1] all around Italy, I estimate a cost function to assess magnitude and significance of costs variation attributable to corruption and mafia presence.

Starting from an original dataset by ANAS, I construct a database containing the necessary variables for each road project. Road works (and their respective cost) are classified by type and activity and then by region and province. Successively, they are assigned indices for road width, cost per metre, geographical variables (terrain ruggedness, population density, rainfall, both regional an province level), and for the variables of interest: mafia and corruption. I exploit several measures to proxy for these latter variables, in particular I use two measures for corruption, extracted from two separate indices: Nifo and Vecchione (2015) *institution quality index* (IQI) and Charron et al. (2019) *Quality of government*, (QoG) one[2]. For organised crime I resort to the mafia presence index (MPI) computed by Transcrime (2013). The unit cost function is derived from a Cobb-Douglas model of cost minimisation and includes controls for environmental conditions, factor prices, work types, and work activities.

The analysis yields 2 major results. First, unespectedly, I don't find evidence of a significant correlation between corruption (and mafia) on unit costs. Although negative, coefficients are small in size and often insignificant. Second, these findings are consistent across several model specifications. They are robust to the addition of regional fixed effects, and refining geographical controls from regional to province level. Moreover, they persist once the sample is restricted to relevant observations like most corrupted regions and provinces, and new

[1] ANAS is one of the major public companies in charge of the national road stock, the other being ASPI

[2] The IQI and the QoG are composite indices, obtained by pondering and aggregating various measures called pillars, to account for overall institutional quality, in the spirit of the World Bank *Worldwide governance indicators* (WGI). Corruption constitutes one of such measures.

road-construction projects. It must be stressed that these findings cannot be considered as causal in the absence of a counterfactual, but rather as statistical associations. However, if my theoretical cost formation framework and project specific controls are correct, they may provide useful insights nonetheless.

Such results are somewhat in contrast with recent empirical literature showing a positive significant relationship between higher expenditures and corruption (Collier et al., 2015; Alexeeva et al., 2011; Olken, 2007)[3] in the road construction sector. It has generally been assumed that bribes firms have to pay to public officials constitute a burden on entrepreneurs, thus rising total production costs Collier et al. (2015). Alternatively, public servant might embezzle part of the funds for public works (Olken, 2007). These aforementioned studies showed indeed a positive variation in costs due to corruption, yet they have been conducted on multinational samples (or in villages within one country) of developing countries.

Interest towards this class of countries has often been led by the shared opinion about their urgent need for infrastructures in their development and how to best allocate funds necessary to their construction. Italy, on the other hand, does not belong to this class. Yet, it presents regional, environmental and social specificities that could make my results (and those from developed countries in general) worthy of further assessment.

The case of Italy. – To begin with, even though Italy has been one of the first nations to adopt new road standards to allow vehicle traffic only, this very primacy has been a double-edge sword for the country. A quickly growing population and a fairly rugged environment made harshly evident how old road projects were not adequate for present traffic. In 2010, Italy had the highest number of private vehicles per person and per km of roads but less than 2000 kilometres of three-lane motorways[4] and freight trucks constituted almost 90% of total goods transported (Marangoni and Marinelli, 2011).

This mix of old and highly trafficked roads levitates upkeep and upgrading costs, whose financial covering has often been neglected. As of today, the network is in charge to two major firms: ASPI, a privately owned company managing the motorway stock, and ANAS, a public firm owning the vast majority of the highways stock and several motorways. In 2017, ANAS estimated a 2,5 billion euro yearly budget for securing and maintaining its highways and motorways stock (about 26 thousands km) while allocating only about 1 billion euros for it. Shockingly, such funds were double the previous years ones (Arona, 2017).

To continue, upgrading works have not been carried out evenly on the territory. In some cases, congested roads have been widened, adding new lanes to fluidify circulation, but this process has been best accomplished in northern Italy where the flat "Padana plain" allowed for easier works. For several of other regions, however, in a largely mountainous and hilly country[5], this has often been impossible, as it would have entailed creating new tunnels and bridges, widening and demolishing existing ones with an unsustainable rise in expenditures. In some cases, infrastructures had to be modified because of insufficient security standards[6],

[3]In Olken (2007) decreased missing funds by up to 27,7% by imposing certainty of ex post auditing on road construction projects

[4]Admittedly, this indicator is an imperfect proxy for traffic since there may be people owning more than a single vehicle, thus raising the cars/km ratio, but it's anyway a first approximation measure

[5]According to the "zone altimetriche" classification by the national bureau of statistics, 76,7% of country's land is either hilly or mountainous

[6]Often these were narrow lanes, short range bends, at-grade crossings etc.

sometimes building ex novo whole road parts, usually new segments bypassing villages they used to cross when traffic volumes were much lower.

Partially as a consequence of such problems, the road stock exhibits a north-south divide that's one of the most evident features of the broader economic fracture between these two macroregions. Both quantitative and qualitative indices highlighted this distinction[7] that public investments have only partially contributed to reduce. As a matter of facts, if stock and time-to-get-to-location indices spot a lack of accessibility for southern regions, indicators of public investments signal a higher expenditure per capita precisely in these areas(Cannari and Franco, 2011). Hence, several studies exploited or tried to assess the link between this discrepancy and corruption (Golden and Picci, 2005; Montanaro, Micucci and Di Giacinto, 2012). While corruption, to various degrees, is present in all of Italy[8], specific forms of organised crime are endemic to some regions and only in more recent years have expanded, first to the country level, then internationally. Since year 2000, 239 municipalities have been dissolved because of mafia infiltration and only 11 of them were located in regions different from Sicily, Calabria, Campania and Puglia. Organised crime undermines economic growth by disrupting market mechanisms that are its fundamental principles in western economies. Legal firms bear direct costs (Rose-Ackerman and Palifka, 2016), such as extortions, as well as indirect ones, like interferences in tendering procedures[9]. These practices may thus create artificial monopolies where firms affiliated to criminal organisations can easily launder money. Cartels, on the other hand, play an ambiguous role. Firms may collude to demand higher prices for tenures, but they can also lead to significant increaes in auction discounts during tendering procedures. Italian anti-corruption authority signaled cases where groups of firms, under coordination of criminal organisations, enacted a scheme of "rotation" whereby each participant was cyclically committed to offering the least expensive bid, thus gaining the tenure (such a scheme has been sometimes referred to as "metodo del tavolino").

Analysing a sample of 30.000 public contracts, Milani et al. (2018) find evidence of such rebates for higher levels of corruption. Conversely, for areas characterized by stronger mafia presence, rebates were much lower than average.

Moreover, firms charging too low prices are more likely to incur in financial losses. Use of inferior quality materials as well as renegotiations of public contracts are sadly not infrequent in Italy. Decarolis (2014) tackles this issue for several Italian regions and finds that half of public administrations' (PAs) savings from first price auctions are lost to following renegotiations and failure in enforcing public contracts. Examining Italian expenditure for public goods provision, Bandiera et al. (2009) show that corruption accounts for only a minor fraction of all money waste. The greater amount thereof is instead attributable to inefficiency

[7] Quantitative indices take into account physical measures such as the value and volume of the infrastructural stock, the population it is supposed to serve, the area where it is distributed and so on. Qualitative indices such as the one from Alampi and Messina (2011) are meant to account for a potential efficient use of infrastructures

[8] ANAC reports that, from 2016 to 2019, Molise and Friuli-Venezia-Giulia were the only regions where there hadn't been arrests linked to corruption in tendering procedures

[9] According to a National anti-corruption authority report, even though money is still the preferred corruptive mean, hiring a family member or a known person is becoming an increasingly popular alternative. (ANAC, 2019). Findings by Scognamiglio (2018) could also lead in this direction, especially in the construction sector.

of public administrations [10].

In this way, income for mafia and corrupted officials might stem more on the extensive margin, rather than on the intensive one: better many medium-size tenures than fewer bulky ones. Findings from Barone and Narciso (2015) and Thomas et al. (2021) seem to point in this direction: through the case of Sicily they show that areas where organised crime is more radicated tend to attract more public investments, even from international institutions such as the EU.

It is clear, then, how the ways corruption and organised crime behave and affect legal activities evolve according to the social and institutional environment they live in, fostering the need for continuous close-look studies.

Literature.– To my knowledge, no other such studies about infrastructures' costs have been carried out considering variations within developed country at its regional level. Many studies have been conducted on the role of corruption on public expenditures (Bandiera et al., 2009; Di Tella and Schargrodsky, 2003; Reinikka and Svensson, 2004; Kenny, 2007; Olken, 2007), and on that of infrastructures on development (Banerjee et al., 2020; Donaldson, 2018; Burgess et al., 2015; Limao and Venables, 2001; Shiferaw et al., 2012; Gertler et al., 2014), yet this paper is more closely related to a narrower strand of literature focussing mainly on the drivers of unit costs for transport infrastructures.

The closest study is Collier et al. (2015), in which the authors exploit a large panel of countries and projects to estimate the impact of corruption and conflict on road construction costs in developing countries. Building on the World Bank ROCKS database (3322 observed projects), they assess the presence of significant differences in unit costs for comparable work activities among countries and sucessively estimate a cost function to quantify the share attributale to corruption, the presence of conflict, and external environmental variables. They show that corruption rises costs on average by 7% when their indicator rises of 1 point over 10 and there is a wedge of 15 percentage points in unit costs between countries whose corruption level is above the median and the rest.

Alexeeva et al. (2008, 2011), on the other hand, carry out two in depth studies about key indicators in the construction sector for two smaller panels of countries, again low and middle-income ones. They find that the number of red flags[11] per country is positively associated to increase in unit costs for their sample units. However, they stress red flags are only a hint of corruption, not a proxy.

[10] The Authors define *waste* the extra cost paid by public bodies for comparable goods and divide it in *active waste* (due to corruption strictu sensu) and *passive waste* (due to administrative inefficiencies). They find that inefficiency share is much larger than that of corruption and that there is no trade-off between these forms of waste. However they highlight that corruption can account for extracosts amounting up to 11% of the good's price. Nevertheless, their analysis is limited to expenditures for standardised goods that amount only to about 40% of Italy's procurement expenditure.

[11] Red flags assigned to contracts are: 1. Period between bid opening and contract signing date is more than 7 months; 2. Cost increases by more than 20 percent during implementation; 3. Time overrun is more than 30 percent of the originally contracted period; 4. Contract value is more than 20 percent above its engineers' estimate; 5. Half or more firms buying bidding documents do not bid; 6. 20 percent or more of prequalified firms do not bid; 7. Difference between winning bid and next lowest bid is within 2 percent; 8. Difference between contract price and read-out bidding price is more than 10 percent; 9. Winning bid is not the lowest bid accepted for detailed examination; 10. Only one or two bidders; and 11. Cost per km for similar work is higher than the 75th percentile.

Although the background of these latter studies is that of development economics, its approach to finding, collecting and classifiying new data would not go amiss in future research concerning already-developed economies.

The rest of the paper is organised as follows: section 2 illustrates the theoretical framework, section 3 describes the data and selection issues, section 4 discusses the econometric results, section 5 provides some future research suggestion and section 6 concludes.

2 Theoretical framework

In the wake of Collier et al. (2015), I employ a regression of work unit costs to assess magnitude and significance of the coefficient associated to the variables of interest (corruption and presence of organised crime).

Work cost can be described by a cost function derived, in this case, from a Cobb-Douglas production function, whose inputs are labour and capital: $Y = A^{-\gamma} l^{\alpha} k^{\beta}$, where Y is the quantity of output produced, A captures inefficiency of production, l and k are labour and capital required, while α and β are their associated coefficients, such that $\alpha + \beta \leq 1$. Thus, the cost minimisation problem reads:

$$\min_{l,k} wl + rk \quad \text{subject to} \quad Y = A^{-\gamma} l^{\alpha} k^{\beta} \tag{1}$$

Let's call C the cost function deriving from the cost minimisation problem, then:

$$C(w,r,Y) = A^{\frac{\gamma}{\alpha+\beta}} w^{\frac{\alpha}{\alpha+\beta}} r^{\frac{\beta}{\alpha+\beta}} Y^{\frac{1}{\alpha+\beta}} \rho \tag{2}$$

Where $\rho = (\frac{\alpha}{\beta})^{\frac{\beta}{\alpha+\beta}} + (\frac{\beta}{\alpha})^{\frac{\alpha}{\alpha+\beta}}$ is a constant. To compare different activities in a sensible way we need a measure of unit cost, such as cost per metre, that's readily obtained by dividing C by Y:

$$\frac{C(w,r,Y)}{Y} = A^{\frac{\gamma}{\alpha+\beta}} w^{\frac{\alpha}{\alpha+\beta}} r^{\frac{\beta}{\alpha+\beta}} Y^{\frac{1-(\alpha+\beta)}{\alpha+\beta}} \rho \tag{3}$$

$\frac{\gamma}{\alpha+\beta}$ is the parameter of interest, whose significance would entails a relevant role of corruption in the cost, if we assume that A is exogenous to the firms. As a matter of facts, the value taken by A is going to be the level of corruption and/or crime in each region as available in our dataset. A higher (lower) level of the parameter implies a shift of the cost function, ceteris paribus rising (decreasing) overall costs.

Taking log of equation (3), adding fixed effects for time, region, work activity and an error term I obtain an estimable equation, that I rewrite as:

$$c_{stu} = \delta \ln A_{st} + \lambda_1 \ln w_{stu} + \lambda_2 \ln r_{stu} + \lambda_3 \ln Y_{stu} + \ln \rho_{stu} + \xi_i + \cdots + \xi_j + \epsilon_{stu} \tag{4}$$

where $c = \ln \frac{C(w,r,Y)}{Y}$, $\delta = \frac{\gamma}{\alpha+\beta}$, w and r are the cost of labour and capital respectively, λ_i their coefficients, ξ is a generic vector of controls and ϵ the regression error term. Pedices s and t identify respectively the region and year of the work, while u is for the work activity.

The parameter δ, then, is going to be the focus of our attention.

Finally, in order to obtain an unbiased estimator, I assume independence in conditional mean for ϵ and X, $E(\epsilon_{stu}|X_{stu}) = 0$, where X_{stu} is the regressor matrix.

There are two major warnings to this model specification.

Even though conditional mean independence is a standard assumption for regression models, in this case it is very restrictive given potential omitted variable bias. I ran several different versions of the model to see whether results change significantly for given specification. It proved not to be the case. Instrumental variables (IVs) are a preferred approach to circumnavigate this issue, even though they shift the causality imputation problem to the validity of the instruments' rather than to that of the initial variables.

A further advantage of IVs is that they allow to infer causality of the regressor even in the absence of a specific control group, the lack thereof in this study prevents one should to be too hasty in regarding my findings as causal. I chose anyway not to use IVs given the exploratory nature of this research.

Reverse causality, on the other hand, should not be an issue. Indeed it's unlikely that geographical controls and project specific variables are influenced by road works unit costs.

Despite these soft spots, the model can still provide valuable insights from its statistical associations.

As I said earlier, there are several ways corruption may affect the cost of road works and not all corruption is of the same kind, or comes from the same source. Constructors may try to bribe public servants in order to be selected for works: in this scenario the bribe would be included in the cost minimisation problem of the firm, leading to higher costs [12]. Alternatively, they might collude in the tendering phase in order to impose preset contract prices. This case has been thoroughly studied by Pesendorfer (2000). As a third option, owners of firms that won the contract could declare greater costs to receive more funds, in order to keep part of them for themselves, or employ lower quality resources with the same intention.

Even though all these alternatives are plausible, the framework I exploit is not suited to specifically detect which one is afoot. Unfortunately I do not possess data about follow-up projects for my sample units, so I cannot even re-estimate an "overall cost" function. I leave to future research, then, questions about the way corruption and mafia affect costs.

3 Data

In order to asses how corruption may affect building costs in the absence of a counterfactual, one needs to capture all its plausible components and isolate the effect of corruption from external confounding factors through an adequate set of controls. Description of such variables and database construction are the content of this section.

The data I use for this study is a sample of 677 road works contracted by ANAS[13],

[12] Actually, inflating contract prices is not the only way firms can incorporate the bribe cost in their operations, but it is not unreasonable to think that at least some part will be included to quickly re-gain to some margin

[13] I had initially been given a set of around 1000 units. The reason for using only this restricted sample thereof is a lack of satisfactory information about all original projects

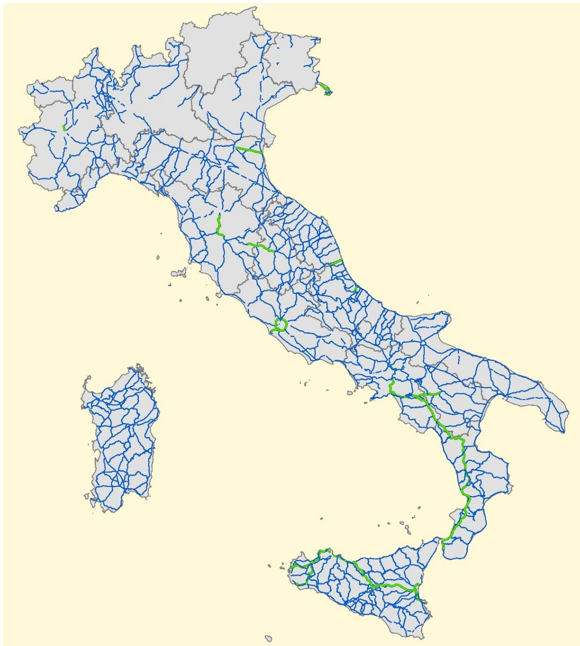

Figure 1: *ANAS road network*

Figure 2: *ASPI road network (Dark green: directly managed. Thin line: contracted out)*

a major government-owned Italian company deputed to construction and maintenance of both highways and motorways. ANAS operates as contracting authority, contracting out works' execution to external firms. Before starting construction, it produces and evaluates a preliminary model. It studies costs, feasibility, and environmental impact of several projects alternatives. Winners of public tenures are generally either those proposing the lowest cost (compatibly with established quality standards) or those whose project meets the highest utility and quality levels (at a pre-established affordable price).

To this day, ANAS manages around 90% of Italian highways and several motorways and motorway junctions, most notably Rome's GRA (grande raccordo anulare) and the notorious A2 "Autostrada del Mediterraneo", whose construction and upgrading process spanned over the last 5 decades.

Despite the bulk of its network, ANAS is not directly in charge of the majority of national motorways, controlled by ASPI (Autostrade per l'Italia), a private company. Therefore motorways are underrepresented in my dataset. This might constitute a first issue with sample's representativeness. I acknowledge the validity of such a problem. However, the choice of the sample at stakes has been conducted taking into account time and availability needed to receive from such firms reliable data that I wouldn't have obtained otherwise. On top of this practical issue, a leading preference factor has been the greater capillarity of the highways system (compared to the motorways one), spreading throughout all Italian regions (except the two autonomous provinces of Trento and Bolzano/Bozen), as well as the diversity of its infrastructures and the need for uniformity in the treatment of data before I received them.

3.1 Description

My first task has been to add necessary information to the sample I had received from ANAS. The initial collection I received was a list of roadworks along with their respective incurred cost, beginning and end date, road name and region they had been executed in. A special office in charge of works on the Salerno Reggio Calabria motorway was counted as a further region by itself. Projects were separated in *new constructions* and *maintenance* works. Maintenance ones were further classified in work types.[14]

To retrieve the data I needed for the analysis, then, I singularly tracked down all sample elements on the national territory. I computed unit costs as cost per metre by simply dividing total cost by the length of the road segment for each project.

Since sample elements were distributed over a 15 years timespan (2004-2019) and works' duration varies according to their type, (from several months for resurfacing to several years for road construction ex novo) it's been crucial to clean costs from the effect of differences in money value over time. To do so, I deflated costs to 2010 euro value, starting from their construction median year. I decided not to use simple CPI as a discount factor, since price trends for consumption good might not correctly reflect those for construction inputs. Thus, I computed a road-infrastructures specific discount rate as an average of the price indices for constructing 1) a gallery, 2) a general road segment, 3) a bridge. Such indices were retrieved from ISTAT's (Italian statistical institute) open data.

Cost per linear metre is simple and convenient measure to compare different activities, however inherits all weak spots from its simplicity. Clearly, differences in road width should increase construction and resurfacing costs (and potentially many others) for each linear metre. So, cost per square metre would have been a more appropriate measure to compare works, since it allows to ignore road width, but it is much harder to compute and would have required to assume an average width per road type so I opted for a different solution and added 2 more variables to capture unit cost variation due to road width. First, a dummy for lanes' number[15], second, another dummy to separate motorways and highways. As a matter of facts, motorways have 3,75 metres wide lanes and a smaller "emergency lane", while highway lanes vary from 3,50 to 3,75 metres in width[16] and have no "emergency lane".

A dummy for projects longer than 10 kilometres has been added too, so as to capture potential economies of scale, that are captured in equation (4) through the term $\ln Y_{stu}$.

Several types of work in the dataset do not concern the carriageway alone: there are barriers upkeep works, plants renovation, bridge and galleries' construction etc. To account for costs coming from different sources I refined the initial classification made by ANAS into a three-level partition. Initially, works are divided in 2 classes: new constructions and maintenance. Subsequently, a finer classification separates them into more specific types such as resurfacing, widening, adding barriers, plants, constructing new segments and so on. Finally, this second level is further refined into peculiar work activities, including the type of added barriers, type of plants, the number of newly construted lanes and the type of new road built.

[14] Work types were: resurfacing, bridges and galleries, barriers, plants, signals.

[15] total road lane number, not just per carriageway

[16] However, they can be as narrow as 2,50 metres per lane for old roads that have not been upgraded to the 3,50 standard, decreed in 2001

Environmental conditions are another important factor to control for. Italy is divided in 20 regions and 107 provinces, that constitute a partition thereof. While all sample works belonged to a single region, some of them spanned across two provinces instead. These were often resurfacing works on a long road segment that crossed 2 provinces[17]. All regions and provinces are reported in the dataset. I also considered the municipalities where activity was carried out, up to a maximum of eight.

It is not hard to imagine why costs for building infrastructures in mountainous areas should on average be superior to those for works on plain ground. Mountainous regions are generally land-locked, far from greater urban centres, and productive facilities: all factors that generally increase transport costs for construction inputs. Changeable meteorological conditions usually require higher security standards for building sites, longer time to complete projects and further difficulties that contribute to levitate costs even more. Moreover, the very type of engeneering solutions adopted in road construction need higher expenditures. Not surprisingly, this variation is accounted for in ANAS' price lists.[18]

Thus, I computed a ruggedness index at regional, provincial, and municipal level for each work.

For all municipalities I know whether their landscape is mainly flat, hilly or mountainous thanks to the ISTAT classification. Such a classification records all municipalities on a discrete scale from 1 to 5, where 1 is mountainous and 5 flat. Exploiting this ranking I associated each work to a ruggedness coefficient at regional, provincial an municipal level. For the regional and provincial ruggedness indices I computed the average of their communes' ruggedness, pondered by their surface. I used the arithmetic mean instead to compute the mountain-presence coefficient at municipal level. The rationale is that the portion of a project crossing a particular municipality is likely independent of its surface area since I assigned to each project all municipalities it crossed, irrespectively of the length corresponding to each of them. Furthermore, had I pondered this latter mean by municipal area, larger communes would have been overrepresented at this level of analysis leading to a biased ruggedness level for infrastructures and providing a less precise estimate for mountain presence along their course.[19]

Meteorological conditions too might affect road construction and upkeep. Angelova and Lupio (2020) provide a full dataset of meteorological variables for 13 EU countries at NUTS3 (provinces for Italy) level for the period 1989 to 2018. Thanks to their information about average monthly precipitations I worked out the yearly values at province and region level.

Population density is another key variable for our estimation purposes. In principle, it can serve as a proxy for traffic on the road at stake. As traffic increases, so does the road's need for upkeep: asphalt is consumed more quickly, galleries need adequate ventilation systems, bridges and barriers have to meet higher safety standards and so on. All these factors are likely to rise costs of construction and maintenance. Another plausible reason, though, is that building and maintaining decisions are taken on the basis of the greatest number of people

[17] Of 48 cross-province works, 33 were resurfacing , 6 barriers, 4 galleries and bridges, 2 plants, 1 upgrading, 1 widening, and 1 new road construction. 45 out 0of 48 were maintenance works.

[18] Construction costs for the same road segment increase by more than a factor of 3 for motorways and extra-urban main roads.

[19] With the same reasoning I computed the simple average of the ruggedness index for works crossing 2 provinces.

they benefit, so that for the same cost, more expensive works are carried out in more densely populated areas. For all these reasons, all sample units are linked to population density (100 people per square km) of the province where they were executed. In the next subsection I will discuss possible biases in the estimation of corruption's impact on road works costs linked to this latter assumption.

I employed GDP from 1982 is selected to proxy for factor prices. I chose 1982 to avoid possible issues with endogeneity. Data for years before 1995 are available only at the macroregional[20] level, so I could include GDP at regional level only from 1995 onwards for additional regressions.

At last, I have got three variables to screen corruption and organised crime, at regional and province level. The first measure is extracted from Nifo and Vecchione (2015) IQI index (institutional quality index). The authors have created the IQI on the basis of the World Bank's WGI (world governance index) aggregating indicators from 5 dimensions to create a single index. Such dimensions are: i) Voice and accountability, ii) Government effectiveness, iii) Regulatory quality, iv) Rule of law, v) Control and corruption. For my study I selected only the coefficient from this last dimension, that's been obtained by further aggregating official data from crimes against public administration, special commissioners and the Golden and Picci (2005) corruption index. This index goes from 0 to 1 where higher values are associated to least corrupted areas and it has been worked out only for years between 2004 and 2019.

Nifo and Vecchione's measure is an extremely valuable one since it is based on official and easily quantifiable data, often hard to obtain and sometimes not completely reliable due to elusion of illicit activities. That's why corruption indices are generally based on surveys and on perceived rather than effective corruption, often reporting that people should be the best detectors of corruptive phenomena as they are directly "on the field"(Charron et al., 2019). For completeness, I also collected data from the European Quality of Government index (QoG)(Charron et al., 2014). The QoG has been created to be the European alternative to the WGI at regional level and as such it collects data from surveys to households and firms. Like with the IQI, I used only the index' dimension concerning corruption, that is normalised between Italian regions and centred on 0 (higher values still correspond to lower corruption). Unfortunately, this measure has been computed and published only four times since it creation and at regional level only, thus I have information only about years 2011, 2013, 2017 and 2021. I linearly interpolated these spot values to obtain values for the gap years. Of course linear evolution of the index is an assumption I can't verify, but it is the simplest (in an Occam's razor perspective) and the most valuable option I had to employ at best the few observations I had.

I resorted to the mafia presence index (IPM) by Transcrime (2013) institute to capture the level of organised crime. This measure is obtained from a synthesis of data up until 2011 about murders, denunciations and dissolved municipalities (and public administrations) due to organised crime, along with information about active organisations and their confiscated properties. Unlike IQI and QoG, IPM takes on higher values for higher mafia presence with no defined upper bound (the lowest and highest value being 0,02 - 101,57 at province level, and 0,31 - 61,21 at regional level.). Since, no updated version of the index has been released

[20]Macroregions are sets of 2 to 6 regions. They are: north-west, north-east, centre, south, and islands.

yet, I rely on this single regional and provincial observation.

As it was predictable, a regression of IPM on IQI identifies a significant positive relationship[21]. So, to avoid collinearity issues, IPM and corruption indices will be studied in separate regressions during the discussion.

Table 1: *Main variables, description and sources*

variable	description	source
Log of cost	Log of unit costs per particular road work activity	ANAS dataset and author's calculations
Ruggedness	Index for presence of mountainous landscape: from -5 (flat) to -1 (mountainous)	ISTAT classification and author's calculations
Log of rainfall	Log of sum of yearly precipitations (2000 - 2018)	Angelova and Lupio (2020)
Population density	Population density (2000 - 2019), 100 people per square kilometre	EUROSTAT
Log of GDP	Log of GDP per capita (1982) per macroregion	ISTAT
# lanes	=1 if the number of lanes is equal to #	author's work
Motorway	=1 if the project involves a motorway	author's work
length $>$ 10 km	=1 if project length is greater than 10 km	author's work
IQI	Index of corruption from Nifo and Vecchione (2015)	Nifo and Vecchione (2015)
log of QoG	Log index of corruption from Charron et al. (2014), updated and shifted to take on only positive values	Charron et al. (2014)
IPM	Mafia presence index by Transcrime institute	Transcrime (2013)

Source: datasets and author's work, as in the table

3.2 Selection

As I previously said, I directly received from ANAS a set of works, but I couldn't exploit all of the observations due to lack of information for some of them. Hence is important to signal

[21]I ran a regression of IPM on IQI (both rescaled so that they go from 1 to 100, from the least to the most corrupted/"mafiosa" region) that evidenced how 1 point increase in corruption is associated to a 1,26 rise in mafia presence, significant at all standard levels.

2 possible biases stemming on the one hand from the selection process I conducted on the basis of available information, on the other from the way data about projects were collected even before they were handed over to me.

1. Assuming that the number of works is proportional to the network's length for each region, an initial χ^2 test (table 12) revealed that such a proportion is not randomly distributed over all regions (significant at all standard levels). Table 12 shows the number of projects for each region and its expected frequency. This value is computed as the product of total works in the dataset I received times the each region's share of the actual national network (total network's dimension is computed as it's total length). It must be said, however, that I could only use present day regional network length to compute expected frequencies for regional works since I had no information about them for previous years. Unfortunately, regional road stock extension has considerably changed over the last 2 decades, so that setting today's regional road network's extension as expected proportion is not the most reliable choice, albeit the only feasible one[22]. Furthermore, many new constructions in the dataset were part of highways that would be ceded to local administrations after their construction. This means they never entered the road stock under ANAS direct management, consequently spoiling the test validity.

   A second χ^2 test (tables 13 and 14), instead, showed that even the amount of projects *included* in the final dataset (remember, they were the ones for which I could find data about their road segment length) is not randomly distributed across regions. In this case I verified whether the overall sample proportion of projects I ended up with were proportional to that of each region. This was not the case. Table 13 divides each region's works between those included in the final dataset (has length) and the rest (no length). Expected frequencies are the product of total regional works I started with times the overall percentage of projects for which I managed to find length. Table 14 repeats the same calculations adding projects from the SA-RC.

   In the light of this preliminary screening, my major concern was that these biases were not due to corruption or mafia themselves, as this would have spoilt the analysis even before the start.

   Discussing with the engineers at ANAS led to the conclusion that the sample contained projects for which they managed to exactly identify cost. I ran 4 further regressions to verify independence between the variables of interest and the amount of sample elements (figures 5 and 6). Corruption and mafia appear to have a sizeable impact only once SARC units are added to the sample. However, there are plausible reasons not to take this result too strictily. As a matter of facts, in the original ANAS sample classification by region these units were not assigned to their respective region; rather, to a separate office. Once added back in, they disproportionally increased their region

[22]During the 90s Italy saw a period of strong decentralisation tendency in public management. This led, among other things, to the transfer of many infrastructures from central administration to local authorities during the early 2000s. In 2018 though, having observed that many local administrations could not satisfactorily manage their infrastructure stock, the "rientro strade" operation was launched, whereby ANAS has gradually been regaining possession of previously owned roads for some time

sample size (Basilicata, Calabria, and Campania). So, Calabria more than doubles its units (from 33 to 80 with assigned length) while Campania adds 21 elements among its projects[23]. Indeed the rise in correlation, as well as its significance, is due to this surge in number of projects for these two regions, that are incidentally those for which corruption and mafia indicators take the highest values (They are represented by the 2 rightmost dots in both graphs in figure 5 and 6). Thus, such relationship is most likely to be due to ANAS' classification methods, that equalized the specialized SARC office to the other regions, rather than to corruption or organised crime.

Moreover, I ran additional regressions to verify potential links between the number of projects for which I found a length measure and corruption and organised crime. As a first guess, one would expect no correlation or even a negative relationship, since my research was based mainly on press releases and public documents about works that criminal organisations should have interest to conceal. This is also what the regressions confirmed (figures 7 and 8). Regressing the share of project included in the final sample for each region on mafia and corruption yielded a weakly positive relationship, not statistically different from zero.

Thus, I take sample units as being selected at least independently of possible corruption-linked episodes.

2. In the dataset one cannot see costs for projects that have not been implemented. This is an important consideration since this truncation in the observed variables produces a bias in the results. Following Collier et al. (2015), the net present value of a project at time t = 0 is: $NPV_0 = -I_0 + (B_1 - C_1)/(1+r) + (B_2 - C_2)/(1+r)^2 + \cdots + (B_T - C_T)/(1+r)^T$, where I_0 and C_t are construction and upkeep costs, while B_t are benefits. So, works selected in the dataset need to have either sufficiently low costs or high enough benefits in order to generate a positive present value. The result is that one ignores projects for which costs are excessive or benefits insufficient. This is the class of projects we do not observe.

   Assume that the model underlying unit costs formation is $c = \beta_0 + \beta_1 x + \epsilon$, where c are costs, β_0 is a constant, x is corruption and ϵ an error term. Let at first β_1 be positive (it is a mainstream assumption that corruption increases construction costs). For any given benefit level, regions with higher corruption levels will be more likely to have excessive costs, fail to generate a positive NPV and so should not appear in the sample. This creates a negative relationship between x and ϵ for observed units, leading to underestimation of corruption's impact. Of course the reverse applies in cases where $\beta_1 < 0$.

Finally, as I quickly mentioned before, greater benefits may depend on the total amount of people interested by the new infrastructure thereby allowing for a positive net present value even in case of higher construction costs if enough people can avoid congestion costs due to lack of infrastructures. Population density has been included as a control variable exactly to account for selection on observables.

[23]The greatest part of the SA-RC motorway lies in these two regions. Out of about 444 total km, 125.7 are in Campania, 289.18 in Calabria, and only 29 in Basilicata. Not surprisingly just 3 projects were added to Basilicata

4 Analysis and discussion

I start by presenting results from equation (4). As a preliminary check, plugging regional fixed effects in and removing time-constant region specific indicators from this equation, I assessed whether the mere fact of being in a different area of the country could lead to a significant rise in project financing. This is of course does not identify corruption, but it's rather a hint for potential results once corruption will be added back in. Surprisingly, almost no regional differences emerged[24].

In the following tables measures for corruption and organised crime are shown parallelly but in different regressions, given the high correlation between these variables. I then introduce a battery of tests and different specifications to assess robustness of the results and to focus on critical subsamples of the dataset. Such tests include adding region fixed effects, reducing geographical control from regional to provincial level and running separate regressions for different work classes and regions.

Table 2 shows an overall not significant impact of corruption and organised crime on costs, that persists even after adding road works from the "Salerno - Reggio Calabria" (SARC) motorway to the sample. Unexpectedly, coefficients are all not significant and small in size[25]. Regressions without SARC units show that unit costs degrease by 0,2% for a 1 point increase in corruption (on a scale from 1 to 100), while the decrease by 0,5% for a unit increase in mafia presence (from 0 to 61,21). At a first glance, such a result is in stark contrast with common knowledge about greater public expenditure for infrastructures in more corrupted areas.

However, there is a some analogy with findings from Milani et al. (2018), who analyse a 30.000 contracts panel and show an increase in discount biddings for more corrupted regions. Conversely, they also find that in zones where mafia organisations have greater influence, discounts are actually smaller, and they explain this difference as the result of the ability of organised crime to group firms into cartels thus managing to charge higher prices. My results, instead, might be due to collusive tactics among firms, often under the regulatory power of established criminal organisations, aiming at lowering bidding offers, but leading to a rise in costs once the works have already begun via the so-called "variazioni in corso d'opera", a sadly common phenomenon in Italy. Unfortunately, I don't possess information about bids and follow-up works in my sample so I cannot group together actvities relating to single major projects and study their behaviour. Explanations about mechanisms through which criminal actors manipulate costs, then, remain only hypotheses grounded in the literature.

Despite this surprising result, overall fit is very high[26] reinforcing confidence in the model. Estimates for unit costs by Collier et al. (2015) display a similar fit, although they apply an

[24]This first screening highlighted few significant differences among regions. Taking Umbria, the least corrupted region, as base category, only Abruzzo showed a significantly different regional coefficient

[25]Furthermore, running the same regression using a dummy for regions whose corruption and/or mafia level is above the median, reveals that for regions with high mafia presence, costs are 25% lower (estimates are significant at 5% level). Coefficients for corruption are of the same magnitude although insignificant. Given this abrupt rise in value, my dummy is probably capturing other regional characteristics that are lessened when using the continuous form of the IPM and IQI indices. Nevertheless, the sign of the coefficient, if not its significance, remains indicative.

[26]the lowest adjusted R^2 is 0,8469

Table 2: *Corruption (IQI) and mafia presence (regional level)*

	Corruption	Mafia presence	Corruption (SARC)	Organised crime (SARC)
IQI	-0,002		0,000	
	(0,004)		(0,004)	
IPM		-0,005		-0,006
		(0,004)		(0,004)
mountainous	0,109	0,114	0,07	0,078
	(0,080)	(0,079)	(0,084)	(0,082)
population density	0,044	0,079	-0,003	0,048
	(0,053)	(0,057)	(0,054)	(0,057)
Log of rainfall	0,063	0,091	0,262	0,323
	(0,228)	(0,229)	(0,228)	(0,233)
Log of GDP1982	0,170	0,064	0,31	0,081
	(0,305)	(0,251)	(0,315)	(0,264)
motorway	0,066	0,092	0,122	0,179
	(0,147)	(0,148)	(0,132)	(0,137)
length > 10 km	-0,173	-0,158	0,082	0,104
	(0,206)	(0,206)	(0,201)	(0,201)
length > 10 km (upkeep)	-1,468***	-1,472***	-1,845***	-1,848***
	(0,237)	(0,237)	(0,233)	(0,233)
3 lanes	1,301	1,244	1,940*	1,95*
	(1,108)	(1,101)	(1,091)	(1,09)
4 lanes	0,541***	0,537***	0,491**	0,486**
	(0,199)	(0,198)	(0,21)	(0,21)
6 lanes	-2,351*	-2,370*	0,011	0,032
	(1,254)	(1,252)	(1,07)	(1,06)
Observations	610	610	676	676
R^2	0,8603	0,8607	0,8469	0,8473

Note: For this and all following tables: dependent variable is the log of cost per metre; all models control for work activity fixed effects, year fixed effects and an interaction between work type and number of lanes fixed effects; *, **, *** denote significance at 10%, 5% and 1% levels.
Source: Author's analysis based on data described in the text.

analogue mehodological framework to a panel of developing countries.

On a whole, geographical controls perform poorly and they too are not significant. Now coefficients are generally positive though, as one would expect for works in harsher conditions or for projects whose construction is likely to benefit a greater mass of citizens.

Road specific variables, on the other hand, seem more promising. Furthermore, they also exhibit a pattern common to almost all model's specifications:

1. Motorways don't appear to require more expensive interventions.
2. There seem to be no evidence of scale economies in the whole sample [27], for works spanning more than 10 kilometres, even though this is not true for maintenance activities whose cost plunges by about 1 time and a half.
3. The "number of lanes" dummies are properly serving their purpose, that is to proxy for road width.

The explanation for apparent scale economies for maintenance activities in table 2 is due to their lower average unit cost, compared to new projects[28], but also to its different behaviour as their length increases.

Indeed, upkeep activities' cost flattens quickly after about 5000 metres (figure 4) and this trend is mainly attributable to resurfacing and barriers-substitution works, that constitute the majority of this subsample as well as the longest work activities. Since such elements in the dataset were almost impossible to track down, I had to rely on information available from their description and length as reported in the original dataset. While I am not aware of any source allowing production gains in road-upkeep, and so entailing such a determinant decrease in unit costs, I am quite confident that this behaviour for maintenace activities is linked to the way their length was registred.

Specifically, after several kilometres of *reported* length, the entity of the projects became probably more spot-like. In this way, as long as the nominal work's length increases, the overall size (in metres) of the work becomes disproportionally large compared to the real activity. This fact is particularly evident for projects' whose length is above 5 km, that is the threshold beyond which unit costs flatten almost completely. Clearly, reported length measures become unreliable for upkeep works at this point. I ran a regression like that in table 2 with new works only in order to verify whether results might change when considering just projects with a sure length measure, but It will be showed in the next section.

As for road width, it's not a coincidence that only 4 lanes roads have on average higher unit costs. Lanes distribution in the sample is not balanced[29] since the vast majority of projects concern 2- and 4-lanes roads. Therefore activities carried out on 3- and 6-lanes roads do not reflect the original sample's work-types' proportions. Actually, they are biased by the unit costs of the most frequent activities on such roads. This is most evident for

[27]I used a dummy for activities longer than 10 kilometres in order to capture a potential decrease in unit costs for bulky works. I added an interaction term between such works and the "maintenance" work class (the other being new constructions) to further refine the analysis of scale economies.

[28]This aspect, however, should largely be captured by my partition of projects into work activities.

[29]In the baseline sample (610 units) lanes distribution is: 68%–2, 2%–3, 28%–4 and 2%–6 lanes. Once the SARC is included they become: 64%–2, 2%–3, 31%–4 and 3%–6 lanes.

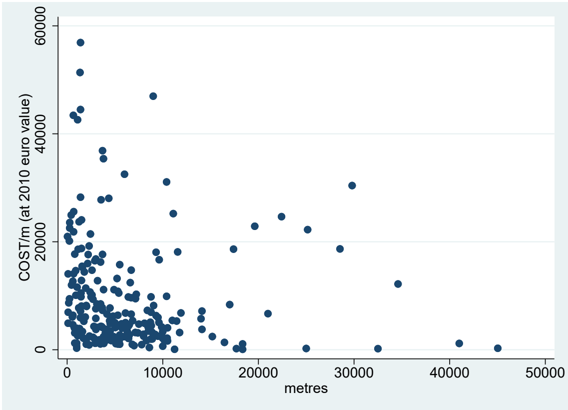

Figure 3: *Unit costs and metres for new constructions*

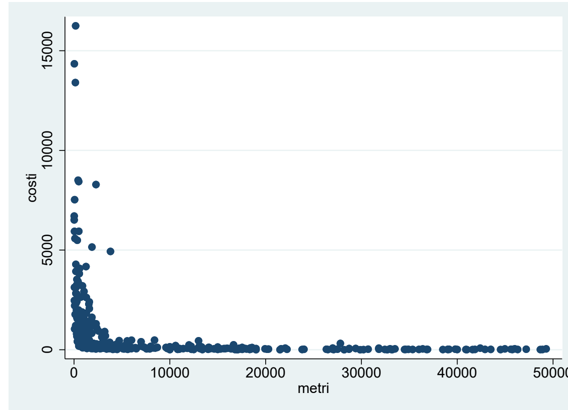

Figure 4: *Unit costs and metres for maintenance works*

projects on 6-lanes roads for which one would reasonably expect higher unit costs. In the baseline dataset, projects on such roads are mainly resurfacing tasks, characterized by low unit costs and as such the coefficient shows a decrease of more than 200%. However, once more expensive activities on the Salerno-Reggio Calabria are included (they are few but have a sizeable impact given the small number of 6 lanes activities in total), the coefficient becomes insignificantly different from zero. Work activity fixed effects clearly cannot capture al variation in unit costs per project, since they are averaged on the whole sample, so the residual variation ends up to be captured by the lanes-number controls. We can see that 3– and 6– lanes controls have indeed large standard errors, a major symptom of collinearity.

In table 3, I turn to corruption as extracted from the quality of government index (QoG) by Charron et al. (2014). Results are not much different from what previously found. Now a 1% increase in corruption lowers unit costs by 0,38%. Nevertheless, estimates for this coefficient are still not significant for most of the models. Exceptionally, once the SA-RC is added to the sample, IMP indicates a significant decrease in costs (albeit only at 10% level). Specifically, unit costs fall by 0,8% following a one point increase in the IPM. Rainfall too registers a significant impact increasing costs by about 0,7% for a 1% rise in precipitations per square kilometre. There are no scale economies for long works, but quite the opposite: greater projects have unit costs roughly 70% higher than smaller ones. For the other road- and activity-specific variables it can be observed the same pattern exposed above. This time, though, 4 lanes works have an even greater costs while 6 lanes ones are less expensive than before. These variations are the consequence of a different distribution of activities within road types, due to the sample truncation necessary to use the QoG index (recall I have QoG index only for years between 2011 and 2019). Using the IQI index on the same sample produces similar estimates in size and sign for all significant coefficients and corruption, while the others, if different in magnitude or direction, are equally not significant and small in size.

4.1 Robustness

Refining the variables of interest and geographical controls to the provincial level does not substantially alter the initial estimates (table 4). Corruption and mafia presence remain

Table 3: *Corruption (QoG) and mafia presence (regional level)*

	Corruption	Mafia presence	Corruption (SARC)	Mafia presence (SARC)
log of QoG	-0,387		-0,296	
	(0,254)		(0,261)	
IPM		-0,006		-0,008*
		(0,005)		(0,005)
Mountainous	-0,036	0,000	-0,095	-0,057
	(0,103)	(0,104)	(0,106)	(0,105)
Population density	0,009	0,046	-0,055	0,011
	(0,055)	(0,073)	(0,055)	(0,072)
Log of rainfall	0,454	0,449	0,647**	0,714**
	(0,292)	(0,314)	(0,276)	(0,279)
Log of GDP1982	-0,309	-0,098	-0,087	-0,14
	(0,380)	(0,315)	(0,374)	(0,321)
Motorway	-0,142	-0,090	-0,024	0,085
	(0,185)	(0,188)	(0,158)	(0,166)
Length > 10 km	0,397	0,500	0,67**	0,76***
	(0,311)	(0,348)	(0,277)	(0,279)
Length > 10 km (upkeep)	-2,027***	-2,157***	-2,44***	-2,556***
	(0,346)	(0,348)	(0,313)	(0,314)
3 lanes	-0,430	-0,757	0,079	-0,15*
	(1,77)	(1,76)	(1,72)	(1,706)
4 lanes	0,837***	0,805***	0,805**	0,777**
	(0,304)	(0,304)	(0,318)	(0,317)
6 lanes	-3,270***	-3,502***	-6,847***	-7,05***
	(0,529)	(0,533)	(1,556)	(1,551)
Observations	412	412	464	464
R^2	0,8481	0,8485	0,8392	0,8399

insignificant and modest in size. Not even a dummy for provinces whose organised crime or corruption levels are above the median suggests a significant role of these variables in manipulating costs. Ruggedness and population density, while rainfall becomes significant only once SARC units are included[30].

As for the baseline specification, physical road type controls continue to show the pattern previously described: maintenance activities have a sizeable lower unit cost, especially on long distances, adding 2 more lanes rises unit costs by about a half while unit costs for works on 3- and 6-lanes highways are very elastic as they are scarcely present in the sample and ill-distributed. I could't run a province level regression for the QoG corruption index since it has not been computed for NUTS 3 areas. Yet, using IQI and IPM indices on the subsample spanning from 2010 to 2019, variables display a behaviour analogous to the regional one exposed in the previous section.

Table 5 illustrates the output of the baseline regression having included regional fixed effects among the regressors. I used provinces as the unit of analysis since all time constant constrols would have been wiped out by fixed effects. As now we have become used to, corruption and organised crime don't appear to incide much on cost variation. I excluded lane variables from the table in order not to repeat myself too often. So, passing to the other controls, this time works on motorways have significant higher cost (30%), that's only present in the SARC sample nonetheless.

4.2 A closer look to regions and work types

In this section I turn to the analysis of a specific subsample of regions and projects.

4.2.1 High corruption and mafia presence

I start by restricting my regional sample to those areas characterised by the highest average values for corruption and mafia presence at both regional and provincial level. Selected regions are those whose average indices are above the third quartile of the distribution. Specifically, I took the mean value of IQI between 2002 and 2019 for each region and province, then I selected those in the fourth quartile. For the IPM index there has been no need of such an aggregation since it already consisted of a single observation per region and province.

I said before that using a dummy when IQI and IPM were above their median level identified a significant decrease in unit costs by about 25%. Thus, the reason for restricting the sample is to check whether for higher levels of corruption and organised crime, costs variations may be more evident than what the previous national level study has revealed[31]. Incidentally, such areas are often from southern Italy whose relevance for this study is twofold.

[30] it is hard to find a rational explanation for this change that is present also in table 3. It might be due to higher precipitation levels for southern regions of Calabria and Campania

[31] It may be that some effects linked to corruption and mafia can be detected only in areas where their presence is sufficiently established. This is what also Milani et al. (2018) report. I accounted for this specificity when using a dummy for regions and provinces whose corruption and mafia presence were above the median. However, this specification is not meant to detect whether corruption plays a more relevant role in more corrupted regions. Rather, a significant change in coefficient would now signal that in more corrupted (and/or "mafiose") regions mafia and corruption produce a higher variation in cost between relatively less and relatively more corrupted (and/or "mafiose") regions

Table 4: *Corruption (IQI) and mafia presence (province level)*

	Corruption	Mafia presence	Corruption (SARC)	Mafia presence (SARC)
IQI	-0,002		-0,001	
	(0,004)		(0,004)	
IPM		-0,002		-0,003
		(0,003)		(0,003)
Mountainous	-0,01	-0,012	-0,011	-0,014
	(0,05)	(0,05)	(0,053)	(0,052)
Population density	0,022	0,023	0,014	0,018
	(0,017)	(0,016)	(0,018)	(0,017)
Log of rainfall	0,233	0,238	0,356*	0,364**
	(0,18)	(0,179)	(0,182)	(0,179)
Log of GDP1982	0,15	0,16	0,229	0,22
	(0,251)	(0,2)	(0,26)	(0,196)
Motorway	0,09	0,10	0,133	0,144
	(0,148)	(0,148)	(0,133)	(0,132)
Length > 10 km	-0,15	-0,15	0,079	0,095
	(0,205)	(0,205)	(0,201)	(0,202)
Length > 10 km (upkeep)	-1,49***	-1,492***	-1,837***	-1,85***
	(0,237)	(0,237)	(0,234)	(0,234)
3 lanes	1,28	1,303	1,968*	1,94*
	(1,102)	(1,101)	(1,093)	(1,090)
4 lanes	0,542***	0,542***	0,483**	0,487**
	(0,198)	(0,198)	(0,210)	(0,210)
6 lanes	-2,46*	-2,437*	0,160	0,084
	(1,256)	(1,254)	(1,062)	(1,06)
Observations	610	610	676	676
R^2	0,8608	0,8608	0,8460	0,8462

Table 5: *Corruption and mafia presence (province level)*

	Fixed effects 1	Fixed effects 2	Fixed effects 3 (SARC)	Fixed effects 4 (SARC)
IQI	0,000		0,003	
	(0,006)		(0,005)	
IPM		0,004		0,001
		(0,004)		(0,003)
Mountainous	-0,094	-0,1	-0,078	-0,08
	(0,064)	(0,065)	(0,068)	(0,068)
Population density	0,02	0,008	0,015	0,017
	(0,019)	(0,020)	(0,019)	(0,019)
Log of rainfall	0,339	0,322	0,416*	0,420*
	(0,209)	(0,208)	(0,219)	(0,219)
Motorway	0,180	0,171	0,295**	0,301**
	(0,161)	(0,161)	(0,146)	(0,147)
Length $>$ 10 km	-0,097	-0,082	0,156	0,145
	(0,208)	(0,208)	(0,204)	(0,204)
Length $>$ 10 km (upkeep)	-1,53***	-1,531***	-1,891***	-1,882***
	(0,24)	(0,24)	(0,234)	(0,234)
Observations	610	610	676	676
R^2	0,8634	0,8623	0,8487	0,84876

First, an immense literature has been produced about the Italian North-South divide[32], a persistent economic gap between such two macroregions. So, it might be of some interest to investigate whether and by how much such divide extends to the road-construction sector. Second, from a more statistical perspective, ANAS operates at national level (with the exeption of Trentino - Alto Adige) but especially in the Mezzogiorno. So my regression could benefit from a more balanced sample where all regions are fairly represented and have a more homogeneous distribution of regional works[33].

Table 6 summarizes results for corruption from the IQI index, both at regional and provincial level, with and without SA-RC units. Sample sizes vary according to the level of analysis. There are 5 regions and 27 provinces respectively[34].

Restricting the sample does not provide much more information than what previous regressions showed. Corruption coefficients are always insignificant, road specific variables exibit a similar behaviour to what has been explained before and geographical controls, apart from rainfall, do not provide further information. A 1% increase in the yearly mean of precipitation is associated with about an equal rise in unit costs for all models in table 6 (excluding the regional level without SA-RC units).

An analogue situation is presented in table 7. Estimates for organised crime presence at regional level are much bigger than they used to be in precedent models. This is due to the variable's lack of variation and reduction of sample size. Indeed, we have only 5 regions and a single value for each of them: this results in high estimates but proportionally high standard errors that prevent us from extrapolating useful information. One can see that for province-level models we are back to usual magnitude standards. Now rainfall is significant and positive for all 4 specifications and shows a similar level, leading to a rise in costs by an amount between 1,06% and 1,21%, according to the model.

4.2.2 New works

Finally, I restrict the whole sample to 4 work *types* belonging to the "new works" class only. These are: i) new road construction, ii) widening, iii) upgrading and iv) new infrastructure construction. In the light of unexpected insignificance of corruption, organised crime and geographic variables so far, there are 2 main reasons to test the model results on this subsample.

1. Observations in this subset have been classified with satisfactory precision and are likely

[32]see for example Cannari and Franco (2011); Daniele et al. (2007); SVIMEZ (2011); Felice (2013); Vecchi (2011). Usually Italy has often been separated in 4 macroregions: northern regions (Valle d'Aosta, Piedmont, Lombardy, Liguria, Trentino-Alto adige, Veneto, Friuli-Venezia giulia, Emilia-Romagna) centre ones (Tuscany, Marche, Umbria and Lazio), south (Abruzzo, Molise, Puglia, Basilicata, Campania, Calabria) and islands (Sicily and Sardinia). When speaking about North-South divide I mean north and centre regions compared to souther ones and islands (this latter group often known as Mezzogiorno), since this is the distinction that's most frequently been used

[33]In the complete sample, regions like Friuli-Venezia giulia and Piedmont are underrepresented with only 6 and 16 observation respectively. Way less than regions whose ANAS network is comparable in size like Val d'Aosta (18 obs) or Puglia (33 obs.)

[34]Regional results here exposed refer to a subsample made of: Calabria, Campania, Molise, Sicily and Sardinia. At province level they refer to: Trieste, Viterbo, Pescara, Taranto, Brindisi and all provinces of Molise, Campania, Calabria, Basilicata and Sicily (except Ragusa)

Table 6: *Corruption (regions and provinces above 3^{rd} quartile)*

	regional	provincial	regional (SARC)	provincial (SARC)
IQI	-0,006	0,000	0,004	-0,002
	(0,01)	(0,007)	(0,010)	(0,007)
Mountainous	-0,291	-0,084	0,040	-0,116
	(0,554)	(0,113)	(0,56)	(0,124)
Population density	0,052	0,017	-0,057	0,020
	(0,119)	(0,03)	(0,115)	(0,031)
Log of rainfall	0,339	0,905**	1,118*	1,055***
	(0,623)	(0,401)	(0,663)	(0,406)
Log of GDP1982	-5,454	0,291	1,58	0,175
	(5,66)	(0,979)	(5,523)	(1,067)
Motorway	0,004	0,035	0,068	0,253
	(0,313)	(0,233)	(0,242)	(0,310)
Length > 10 km	-0,45	0,162	0,248	0,260
	(0,319)	(0,355)	(0,305)	(0,310)
Length > 10 km (upkeep)	-1,367***	-1,677***	-2,049***	-1,978***
	(0,375)	(0,41)	(0,375)	(0,372)
Observations	214	235	278	300
R^2	0,8599	0,841	0,8161	0,8129

Table 7: *Mafia presence (regions and provinces above 3^{rd} quartile)*

	regional	provincial	regional (SARC)	provincial (SARC)
IPM	-0,184	0,000	0,114	-0,002
	(0,155)	(0,004)	(0,146)	(0,003)
Mountainous	2,964	-0,057	-2,440	-0,081
	(2,76)	(0,148)	(2,585)	(0,148)
Population density	1,575	0,019	-0,913	0,023
	(1,273)	(0,027)	(1,185)	(0,029)
Log of rainfall	1,064**	1,062***	1,134**	1,215***
	(0,549)	(0,363)	(0,552)	(0,363)
Log of GDP1982	-13,627	-0,357	7,199	0,181
	(10,468)	(0,601)	(9,728)	(0,521)
Motorway	0,020	-0,175	0,272	0,146
	(0,287)	(0,289)	(0,243)	(0,234)
Length > 10 km	0,162	0,412	0,679**	0,556
	(0,334)	(0,399)	(0,317)	(0,345)
Length > 10 km (upkeep)	-1,772***	-1,889***	-2,23***	-2,236***
	(0,397)	(0,446)	(0,376)	(0,403)
Observations	218	205	282	269
R^2	0,8705	0,8747	0,8365	0,8339

to be affected by geographical controls. I managed to retrieve sufficient information about these projects via a time-consuming search of ANAS press releases and project descriptions for nearly all of them. These works needed on average longer time than the rest of the sample and were the most expensive ones. While they had a median length of 4133 metres, only about 1,4 kilometre less than maintenance works, their mean unit costs is 8297 euro per linear metre against 760 for upkeep activities, on average.

Therefore, two hypotheses follow. On the one hand, by looking at figure 3, the light negative relationship between length and cost is an encouraging signal about the absence of excessively registered length compared to real length. On the other, if I can reasonably hope that "intensity" of the work was quite uniform on their total length (works should be less spot-like),they should be more likely affected by geographical factors such as terrain ruggedness, adverse weather conditions and crossing densely populated areas. Indeed such factors may have lower influence on small-size activities or could be easily captured by work activitiy fixed effects. I can indirectly verify whether this hypothesis holds by looking at significant changes in the coefficients for environmental controls. It's not hard to imagine that costs for gallery plants or road resurfacing are going to be similar across regions irrespectively of how mountainous these are, as long as this does not imply significantly higher transport costs for production inputs. Conversely, widening or constructing a new road segment will probably be affected by mountainous elevations, valleys, cities to cross, as well as meteorological conditions during construction.

2. Such projects can be both more attractive for criminal organisations and sensitive to corruption given the greater volume of payments involved. It could be esier for public servants and firms to hide illegal collateral transactions within funding operations for these great projects. Furthermore, money laundering may benefit an acceleration in its process thanks to bigger contracts.

Table 8 and 9 summarize results for corruption and mafia presence models at regional and province level respectively. As expected, now geographical controls play a significant role. In table 8, an increase of 1 point on a 4 points scale for mountainous elevations, raises units cost by 33% in all specifications with maximum significance. A 4 point increase (that is passing from flat to mountainous environment) would entail a rise by 132%. Population density too suggests that for every 100 more people per squared mile, costs rise by 12% at least. This coeffecent loses significance only in the 3^{rd} specification (corruption including SARC observations). Consistently to what already found so far, no scale economies are present. Motorways works, instead, now appear to require higher unit costs. Nevertheless, it's probably going to be a simple statistical association given it's slight significance and lack thereof once SA-RC observations are pooled in.

Corruption and mafia are still modest in size and insignificant, except for the case of IPM with the SA-RC. Yet, significance is only at 10% level. Notably, adjusted R^2 is now much lower than for the previous models, even for ones comparable in sample size[35]. As displayed in figure 3, new works exhibit much more variation in unit costs compared to upkeep ones.

[35]see table 7 and 6

Table 8: *Corruption and mafia presence (regional level, new works only)*

	Corruption	Mafia presence	Corruption (SARC)	Mafia presence (SARC)
Corruption (IQI)	-0,002		-0,009	
	(0,006)		(0,006)	
Mafia presence (IPM)		-0,007		-0,011*
		(0,006)		(0,006)
Mountainous	0,334***	0,35***	0,338***	0,340***
	(0,113)	(0,113)	(0,130)	(0,129)
Population density	0,127*	0,172**	0,122	0,150*
	(0,076)	(0,08)	(0,083)	(0,086)
Log of rainfall	-0,32	-0,241	-0,349	-0,236
	(0,446)	(0,34)	(0,376)	(0,384)
Log of GDP1982	0,139	-0,105	-0,012	-0,071
	(0,446)	(0,397)	(0,483)	(0,444)
Motorway	0,461*	0,484*	0,353	0,406
	(0,259)	(0,256)	(0,262)	(0,262)
Length > 10 km	-0,227	-0,203	0,114	0,164
	(0,192)	(0,192)	(0,201)	(0,202)
Observations	199	199	227	227
R^2	0,5691	0,573	0,4621	0,4649

Table 9: *Corruption and mafia presence (province level new works only)*

	Corruption	Mafia presence	Corruption (SARC)	Mafia presence (SARC)
Corruption (IQI)	-0,006		-0,01*	
	(0,006)		(0,006)	
Mafia presence (IPM)		-0,006		-0,001
		(0,004)		(0,003)
Mountainous	0,145**	0,148**	0,13*	0,130*
	(0,07)	(0,07)	(0,08)	(0,080)
Population density	0,077***	0,079***	0,86***	0,066**
	(0,029)	(0,027)	(0,031)	(0,030)
Log of rainfall	0,102	0,13	0,112	0,007
	(0,26)	(0,259)	(0,288)	(0,286)
Log of GDP1982	-0,098	-0,095	-0,180	-0,237
	(0,367)	(0,31)	(0,402)	(0,327)
Motorway	0,367	0,304	0,280	0,217
	(0,262)	(0,26)	(0,263)	(0,262)
Length > 10 km	-0,14	-0,16	0,186	0,178
	(0,192)	(0,191)	(0,202)	(0,204)
Observations	199	199	227	227
R^2	0,5709	0,5743	0,4648	0,4571

Therefore, restricting the analysis to this class only and reducing the number of activitiy fixed effect from 39 to 15 is responsible for the rise in unexplained variation for this specification[36].

Results from table 9 are not much different, though one can see how ruggedness' impact has sensitively decreased in size. Population density displays the same trend showing a rise by 7% in costs every 100 people per square kilometre. Its effect, though, is now significant even at 99% conficence level. Motorways have no different unit costs than highways and corruption, scale economies, and R^2 are in line with previous results.

5 Future research

Contrarily to what often hypothesized, this paper showed that corruption and organised crime have no impact on cost for construction and maintenance of the Italian road network. Yet, this study still needs further analyses to complement its weaknesses and to reinforce the stance of such an unexpected result.

First of all, little less than seven hundred observations are surely not enough to represent the whole road system. Collaboration with public institutions such as ANAS should therefore be aimed at increasing sample size and road works description so as to benefit following research about the topic. In creating my small dataset for transport infrastructures, I took inspiration from the World Bank ROCKS database, but there is likely no doubt that its construction principles could be productively extended to different topics. Nevertheless, a specialized pool of engineers and professionals is going to be needed to expand sample size and provide useful works' description and categorizations. Highlighting the role of cooperation between public institutions and research poles in providing adequate information. Indeed, my initial χ^2 tests assessed that corruption and mafia are not the reasons behind sample selection. Nevertheless, a formal collection and distribution process for such specific data is a necessary condition for a complete, thorough analysis.

Second, deeper investigations in the role of corruption and mafia on unit costs would be asked not only to expand their sample base, but also to provide a causal link between costs and the variables of interests that this study was has not been able to explain. Instrumental variables are generally the preferred way when dealing with standard regressions as in this paper. Barone and Narciso (2015) and Thomas et al. (2021) show an increase in the allocation of public funds in areas where organised crime is more intense. To instrument for mafia presence they rely on two main sources: 1) data about land productivity shifters like terrain slope, altitude, and rainfall between 1850-1860, that are showed to affect mafia location back then, given its agricultural roots, but are uncorrelated with it nowadays, 2) the Damiani-Jacini parliamentary enquiry from 1885, measuring historical roots of mafia in several areas. Such an approach appears a natural extension to this study.

Moreover, in the framework of this paper I have been almost completely agnostic about the way corruption and mafia act in practice on prices for public tenures. Variations in unit costs were supposed to stem either from bribes paid to public officials, or from cartels among firms

[36]Fixed effects capture less variation in this specification. Two main differences with the initial sample are worth noting: i) observations have decreased proportionally more than the number of activities, ii) fewer work activities account for the majority of works. In particular, observations shrunk from 610 to 199 as opposed to 39 and 15 for work activities and the 4 most frequent activities accout for 151 works.

in order to demand higher payments, as well as from agreements to charge the lowest bid, win the tenure and cyclically share public contracts. Such trends can largely vary between areas and display variation over time. With this regard, quantitative studies at aggregate level are complementary to case-specific analyses since they allow to easily associate social conditions (such as mafia presence) and economic outcome at macro level. Nevertheless, there is a research gap to fill in determining the way the former interacts with the latter without narrowing the analysis to a case-specific study. This is an important lacuna for a country like Italy where corruption and organised crime have and endemic and lastly presence, especially in the construction sector. As an example, Porter and Zona (1993) provide a test to detect cartels aiming at manipulating estimates of state highways construction in the U.S. Pesendorfer (2000) exploits biddings for the milk provision service in Texas and Florida and shows two distinct ways local cartels operated[37].

Finally, even if there are no precise estimations, many transport infrastructures suffer from construction and maintenance with poor quality materials or even lack thereof[38]. This allows constructors to curb costs while loading a double burden on society: infrastructures users face a safety risk in their activities, while works financers are likely to incur subsequent expenditures in further upkeep. Thanks to monopolistic positions in inputs' production (cement, contcrete etc.) and specialised building firms, criminal organisations can safely invest in the construction sector where they have control over the whole productive chain Investments in the construction sector are appealing for criminal organisations as they can have assuring control over a network that goes from input and production to sale (Sciarrone, 2021; Varese, 2011). An intersectoral analysis of unit costs for different kinds of projects would be a fruitful avenue for future research.

6 Conclusion

The purpose of this paper was to detect the impact of mafia and corruption on unit costs of public road infrastructures in Italy. The country provided a sensible background for such a study given its variation in the endemic nature of organised crime, corruption levels and social-geographical differences between regions.

Estimating a cost function for my sample of road works didn't generally yield a significant effect of mafia and corruption and, contrarily to what generally supposed, it sometimes showed a significant negative effect. Such counterintutitive results deserve a closer investigation both at aggregate and case level, concerning the way criminal organisations and social phenomena like corruption operate in the national and regional environment. Perverse anti-competitive behaviours like the so called "metodo del tavolino" and collusive tactics are strategies that vary according to regional specificities as well as introduction of new regulations[39]

[37]Cartels in these two states had different fashion in allocating contracts: one divided the market among members, the other bribed competitors to refrain from bidding. The author outlines a theoretical framework for this phenomenon first, then proceeds to test it with available data.

[38]This may well occur when firms have gained public tenures at too low prices as a consequence of collusion, all the more.

[39]Parliament's decree n. 50, year 2016, for example, introduces incentives to new tenuring forms exactly to avoid mere Dutch biddings and to curb anticompetitive strategies.

Furthermore, the very sampling and classification process proved to be much harder than expected for a developed country and posed imputation problems that might have partially biased my results[40]. Retrieving adequate information about contracts and data about public tenures should be granted more easily by efficient institutions. This cannot be the burden of research institutions alone as it requires professional skills in quantifying features about contracts and carried out works.

[40]I spotted the major issue regarding the nature of upkeep interventions at the beginning of the "analysis and discussion" paragraph.

References

Alampi, Demetrio and Giovanna Messina, "Time-is-money: i tempi di trasporto come strumento per misurare la dotazione di infrastrutture in Italia," *Le infrastrutture in Italia: dotazione, programmazione, realizzazione*, 2011, pp. 137–174.

Alexeeva, Victoria, Cesar Queiroz, and Satoshi Ishihara, "Monitoring road works contracts and unit costs for enhanced governance in Europe and Central Asia," 2011, *transport paper 33.*

_ , Gouthami Padam, and Cesar Queiroz, "Monitoring road works contracts and unit costs for enhanced governance in Sub-Saharan Africa," 2008, *transport paper 21.*

ANAC, *La corruzione in Italia 2016-2019. Numeri, luoghi e contropartite del malaffare*, Autorità nazionale anticorruzione, Italy, 2019.

Angelova, Denitsa and Norman Blanco Lupio, "Constructing a meteorological indicator dataset for selected European NUTS 3 regions," *Data in brief*, 2020, *31*, 105786.

Arona, Alessandro, "Per la manutenzione delle strade servono 2,5 miliardi l'anno," *il sole 24 ore*, 2017.

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

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the road: Access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, 2020, *145*, 1024–42.

Bardhan, Pranab, "Corruption and development: a review of issues," *Journal of economic literature*, 1997, *35* (3), 1320–1346.

Barone, Guglielmo and Gaia Narciso, "Organized crime and business subsidies: Where does the money go?," *Journal of Urban Economics*, 2015, *86*, 98–110.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel, "The value of democracy: evidence from road building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–51.

Cannari, Luigi and Daniele Franco, "Il Mezzogiorno: ritardi, qualità dei servizi pubblici, politiche," *Stato e mercato*, 2011, *31* (1), 3–40.

Charron, Nicholas, Lewis Dijkstra, and Victor Lapuente, "Regional Governance Matters: Quality of Government within European Union Member States," *Regional Studies*, 2014, *48* (1), 68–90.

_ , Victor Lapuente, and Paola Annoni, "Measuring quality of government in EU regions across space and time," *Papers in Regional Science*, 2019, *98* (5), 1925–1953.

Collier, Paul, Martina Kirchberger, and Måns Söderbom, *The cost of road infrastructure in low and middle income countries*, The World Bank, 2015.

Daniele, Vittorio, Paolo Malanima et al., "Il prodotto delle regioni e il divario Nord-Sud in Italia (1861-2004)," *Rivista di politica economica*, 2007, *97* (2), 267–316.

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

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, April 2018, *108* (4-5), 899–934.

Felice, Emanuele, *Perché il Sud è rimasto indietro*, il Mulino Bologna, 2013.

Gertler, Paul J, Marco Gonzalez-Navarro, Tadeja Gracner, and Alexander D Rothenberg, "The role of road quality investments on economic activity and welfare: Evidence from Indonesia's highways," *Unpublished manuscript*, 2014.

Golden, Miriam A and Lucio Picci, "Proposal for a new measure of corruption, illustrated with Italian data," *Economics & Politics*, 2005, *17* (1), 37–75.

Kenny, Charles, "Construction, corruption, and developing countries," *World Bank Policy Research Working Paper*, 2007, (4271).

Limao, Nuno and Anthony J Venables, "Infrastructure, geographical disadvantage, transport costs, and trade," *The world bank economic review*, 2001, *15* (3), 451–479.

Marangoni, Daniele and Giuseppe Marinelli, "Il crescente ruolo delle Amministrazioni locali nella programmazione e gestione della viabilità stradale," *Banca d'Italia, Le infrastrutture in Italia: dotazione, programmazione, realizzazione, Seminari e Convegni*, 2011, *7*, 619–648.

Mauro, Paolo, "Corruption and growth," *The quarterly journal of economics*, 1995, *110* (3), 681–712.

Milani, Riccardo, Francesco Calderoni, Carlotta Carbone, and Martina Rotondi, "L'impatto di corruzione e mafia sugli appalti pubblici: un'esplorazione empirica," 2018.

Mo, Pak Hung, "Corruption and economic growth," *Journal of comparative economics*, 2001, *29* (1), 66–79.

Montanaro, Pasqualino, Giacinto Micucci, and Valter Di Giacinto, "L'impatto macroeconomico delle infrastrutture: una rassegna della letteratura e un'analisi empirica per l'Italia," *L'impatto macroeconomico delle infrastrutture: una rassegna della letteratura e un'analisi empirica per l'Italia*, 2012, pp. 19–59.

Nifo, Annamaria and Gaetano Vecchione, "Measuring institutional quality in Italy," *Rivista economica del Mezzogiorno*, 2015, *29* (1-2), 157–182.

Olken, Benjamin A, “Monitoring corruption: evidence from a field experiment in Indonesia,” *Journal of political Economy*, 2007, *115* (2), 200–249.

Pellegrini, Lorenzo and Reyer Gerlagh, “Corruption’s effect on growth and its transmission channels,” *Kyklos*, 2004, *57* (3), 429–456.

Pesendorfer, Martin, “A study of collusion in first-price auctions,” *The Review of Economic Studies*, 2000, *67* (3), 381–411.

Porter, Rob and John Zona, “Detection of Bid Rigging in Procurement Auction,” *Journal of Political Economy*, 06 1993, *101*, 518–38.

Reinikka, Ritva and Jakob Svensson, “Local capture: evidence from a central government transfer program in Uganda,” *The quarterly journal of economics*, 2004, *119* (2), 679–705.

Rose-Ackerman, Susan and Bonnie J Palifka, *Corruption and government: Causes, consequences, and reform*, Cambridge university press, 2016.

Sciarrone, Rocco, *Mafie vecchie, mafie nuove: radicamento ed espansione*, Donzelli editore, 2021.

Scognamiglio, Annalisa, “When the mafia comes to town,” *European Journal of Political Economy*, 2018, *55*, 573–590.

Shiferaw, Admasu, Måns Söderbom, Eyerusalem Siba, and Getnet Alemu, “Road networks and enterprise performance in ethiopia: Evidence from the road sector development program,” *International Growth Centre, The College of William and Mary, Working Paper*, 2012, *12*, 0696.

SVIMEZ, *150 anni di statistiche italiane: Nord e Sud: 1861-2011*, Il Mulino, 2011.

Tella, Rafael Di and Ernesto Schargrodsky, “The role of wages and auditing during a crackdown on corruption in the city of Buenos Aires,” *The Journal of Law and Economics*, 2003, *46* (1), 269–292.

Thomas, Marlene, João Pereira dos Santos, and José Tavares, “Does Organized Crime Attract EU Funding? Evidence from the Sicilian Mafia,” 2021.

Transcrime, “Progetto PON Sicurezza 2007–2013: Gli Investimenti Delle Mafie. Rapporto Linea 1,” 2013.

Varese, Federico, *Mafias on the Move*, Princeton University Press, 2011.

Vecchi, Giovanni, *In ricchezza e in povertà: il benessere degli italiani dall’Unità a oggi*, Il mulino, 2011.

7 Appendix

Table 10: 3-levels road works partition

Work class	Work type	Work activity	
New work	upgrading	securing	5
208	18	junction	2
		improved quality	4
		undefined	7
	widening	+ 2 lanes	32
	48	new road section	15
		undefined	1
	new road construction	junction	8
	113	new road	40
		bypass road	64
		auxiliary road	1
	barriers	metal divider	2
	4	generic barriers	2
	plants	technological plants	3
	3		
	bridges and galleries (B&G)	strengthenin gallery	1
	21	strengthenin bridge	4
		restructuring bridge	1
		new gallery	6
		new bridge	8
		retructuring gallery	1
	surface	resurfacing	1
	1		
Maintenance	barriers	anti noise barriers	4
403	76	generic	39
		anti rocks b	11
		bridge	14
		cement divider	5
		metal divider	3
	resurfacing	securing	1
	197	strengthening	26

Continued on next page

Table 10 – *Continued from previous page*

Work class	Work type	Work activity	
		resufarcing	85
		refurbishment	46
		surface (roadside objects included)	39
	plants	fire plants	5
	47	light	9
		light + gallery upkeep	2
		upkeep damaged plants	2
		light and ventilation	2
		light, ventilation & security	2
		technological	24
		undefined	1
	bridges and galleries (B&G)	road widening	1
	79	strengthening bridge	9
		expansion joints	14
		securing	3
		exp joints & barriers	4
		refurbishing decks	8
		refurbishing slab & resurfacing	1
		refurbishing slab	9
		restructuring bridge	30
	signals	upgrading signals	2
	4	new signals	2

Table 11: regional road–network extension

Region	highways (km)	motorways (km)	total (km)	% km per region
abruzzo	1.432,72	14,7	1.447,42	0,05
basilicata	983,62	80,484	1.064,10	0,04
calabria	1.685,95	294,801	1.980,75	0,07
campania	1.325,67	160,462	1.486,13	0,06
emilia romagna	1.866,73	49,282	1.916,02	0,07
friuli	166,646	33,027	199,67	0,01
lazio	1.142,57	85,586	1.228,16	0,05
liguria	635,943	0	635,94	0,02
lombardia	1.981,71	0	1.981,71	0,07
marche	1.236,10	26,3	1.262,40	0,05
molise	667,002	0	667,00	0,02
piemonte	1.596,82	10,65	1.607,47	0,06
puglia	1.492,60	0	1.492,60	0,06
sardegna	2.944,36	0	2.944,36	0,11
sicilia	3.492,24	424,085	3.916,33	0,15
toscana	1.412,95	75,45	1.488,40	0,06
umbria	769,356	39,38	808,74	0,03
val d'aosta	143,612	0	143,61	0,01
veneto	735,381	0	735,38	0,03
TOTAL	25.711,98	1.294,21	27.006,19	1,00

Table 12: number of works perregion and expected frequencies

Region	regional works	regional works (SARC)	EXP n° works	EXP n° works (SARC)
abruzzo	58	58	51	55
basilicata	66	68	37	40
calabria	45	97	70	75
campania	54	75	52	56
emilia romagna	63	63	67	73
friuli	10	10	7	8
lazio	64	64	43	46
liguria	20	20	22	24
lombardia	57	57	70	75
marche	32	32	44	48
molise	36	36	23	25
piemonte	56	56	56	61
puglia	46	46	52	56
sardegna	97	97	103	111
sicilia	90	90	138	148
toscana	44	44	52	56
umbria	49	49	28	31
val d'aosta	19	19	5	5
veneto	43	43	26	28
TOTAL	949	1024		

The χ^2 test p-value for standard and SARC sample is, respectively: 6,847E-21 and 1,465E-20

Table 13: observed and expected frequencies for projects per region (non SARC units)

regions	projects	has length	no length	has length (expected)	no length (expected)
Abruzzo	58	36	22	38	20
Basilicata	66	50	16	43	23
Calabria	45	33	12	30	15
Campania	54	36	18	35	19
Emilia Romagna	63	40	23	41	22
Friuli	10	7	3	7	3
Lazio	64	46	18	42	22
Liguria	20	13	7	13	7
Lombardy	57	37	20	37	20
Marche	32	25	7	21	11
Molise	36	26	10	24	12
Piedmont	56	17	39	37	19
Puglia	46	33	13	30	16
Sardinia	97	50	47	64	33
Sicily	90	72	18	59	31
Toscana	44	29	15	29	15
Umbria	49	30	19	32	17
Val d'aosta	19	18	1	12	7
Veneto	43	25	18	28	15

Overall sample frequencies for projects with and without associated length are, respectively: 0,66 and 0,34

The χ^2 test p-value is 2,25E-7

Table 14: observed and expected frequencies for projects per region (SARC units)

Regions	projects	has length	no length	has length (expected)	no length (expected)
Abruzzo	58	36	22	39	19
Basilicata	68	52	16	46	22
Calabria	97	80	17	66	31
Campania	75	57	18	51	24
Emilia Romagna	63	40	23	43	20
Friuli Venezia Giulia	10	7	3	7	3
Lazio	64	46	18	43	21
Liguria	20	13	7	14	6
Lombardy	57	37	20	39	18
Marche	32	25	7	22	10
Molise	36	26	10	24	12
Piedmont	56	17	39	38	18
Puglia	46	33	13	31	15
Sardinia	97	50	47	66	31
Sicily	90	72	18	61	29
Toscana	44	29	15	30	14
Umbria	49	30	19	33	16
Valle d'Aosta	19	18	1	13	6
Veneto	43	25	18	29	14

Overall sample frequencies for projects with and without associated length are, respectively: 0,68 and 0,32

The χ^2 test p-value is 4,682E-10

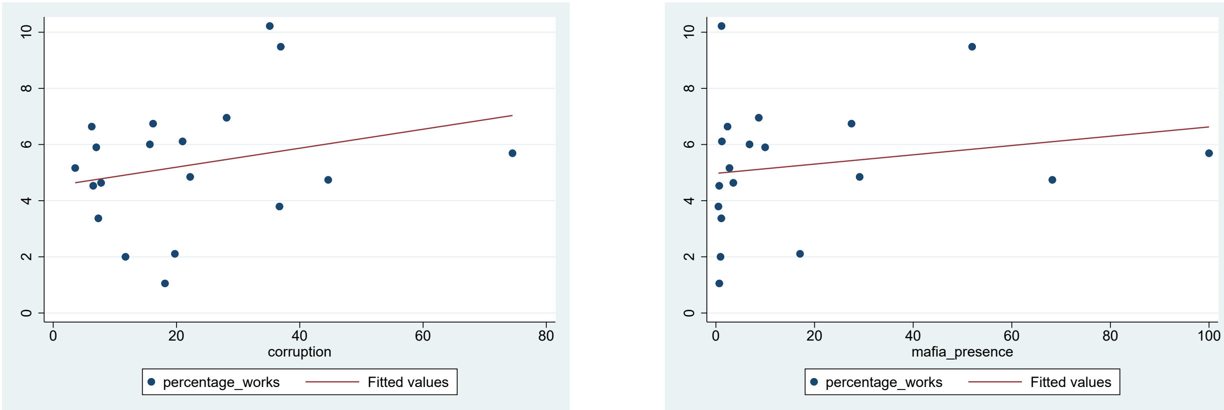

Figure 5: *Correlation beween projects per region (as a percentage of the whole sample, SARC elements excluded) and average IQI corruption and IPM. Coefficients are not significant at all standard levels.*

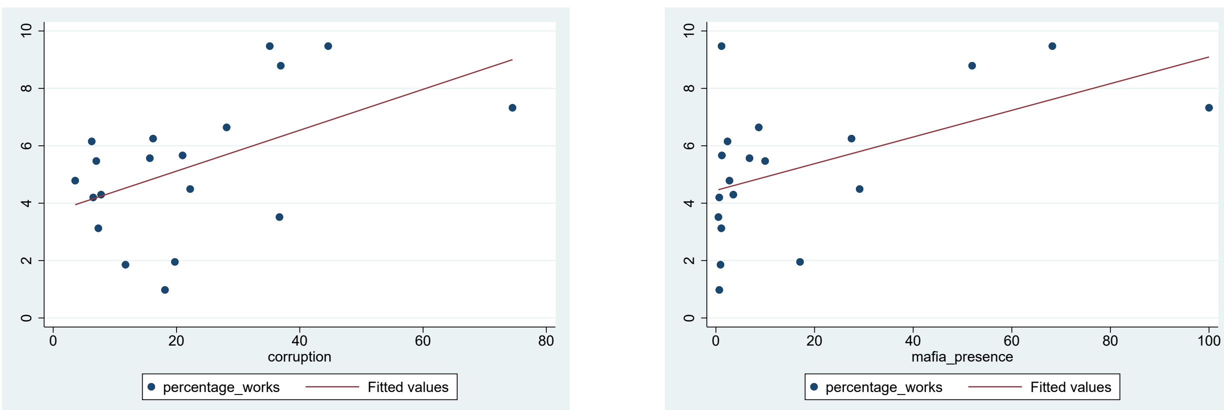

Figure 6: *Correlation beween projects per region (as a percentage of the whole sample, SARC elements included) and average IQI corruption and IPM. Coefficients are significantat 5% level*

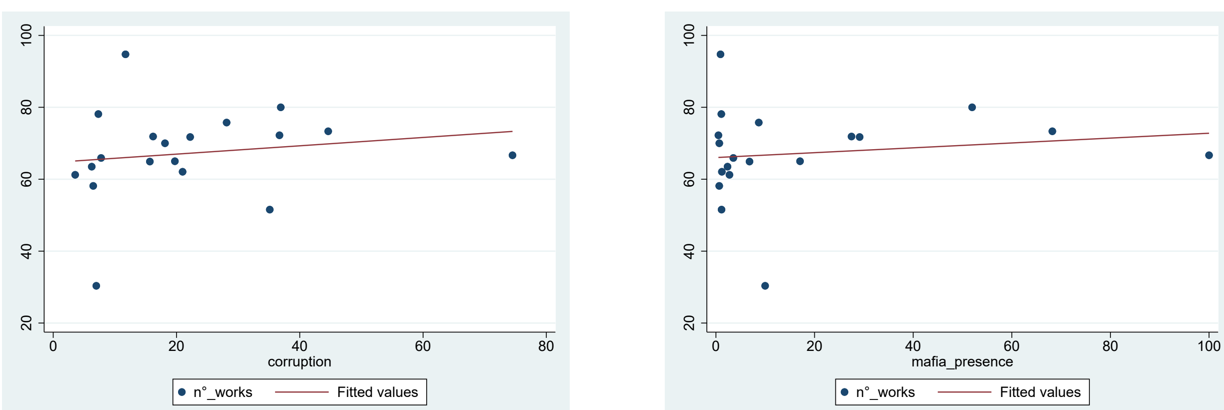

Figure 7: *Correlation beween each region's share of projects included in the sample (SARC elements excluded) and average IQI corruption and IPM.*

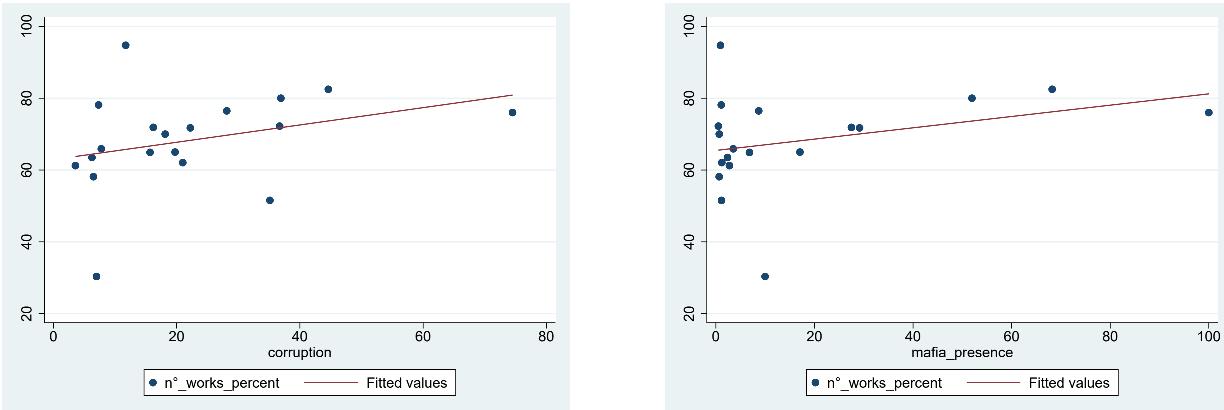

Figure 8: *Correlation beween each region's share of projects included in the sample (SARC elements included) and average IQI corruption and IPM.*

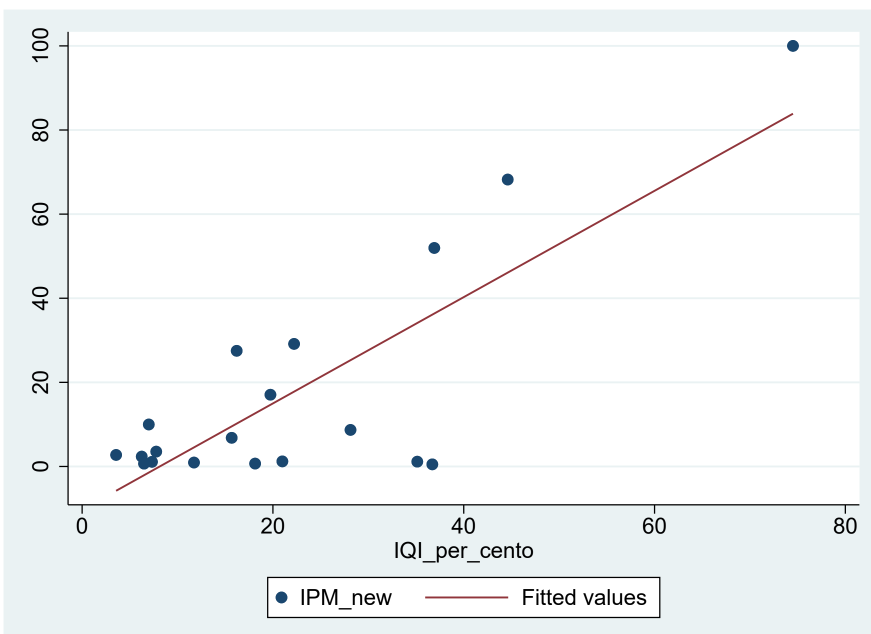

Figure 9: *Correlation between IPM and IQI corruption*