Pirate Attacks and the Shape of the Italian Urban System

Antonio Accetturo, Michele Cascarano

and Guido de Blasio

UNIVERSITÀ DEGLI STUDI DI TRENTO

Dipartimento di Economia e Management

Università degli Studi di Trento

Department of Economics and Management, University of Trento, Italy.

Editors

Guidelines for authors
Papers may be written in Italian or in English. Faculty members of the Department must submit to one of the editors in pdf format. Management papers should be submitted to R. Gabriele. Economics Papers should be submitted to L. Andreozzi. External members should indicate an internal faculty member that acts as a referee of the paper.

Typesetting rules:
1. papers must contain a first page with title, authors with emails and affiliations, abstract, keywords and codes. Page numbering starts from the first page;
2. a template is available upon request from the managing editors.

PIRATE ATTACKS AND THE SHAPE OF THE ITALIAN URBAN SYSTEM[•]

Antonio Accetturo
Bank of Italy

Michele Cascarano
University of Trento and Bank of Italy

Guido de Blasio
Bank of Italy

August, 2019

Abstract
From the sixteenth to the early nineteenth century, coastal areas of Italy (especially, in the south-west) were subject to attacks by pirates launched from the shores of Northern Africa. This paper documents that, in order to protect themselves, residents of coastal locations moved inland to mountainous and rugged areas. It also shows that such relocation constrained local economic development for a long period after the piracy threat had subsided. By hampering the growth of major urban centers, the attacks may have also had aggregate consequences on Italy's post-WWII development.
JEL Classification: R1, N9, O1
Keywords: City size distribution; Historical shocks; Local development; Aggregate effects

1. Introduction

The spatial distribution of population within a country is far from being homogenous. All countries are characterized by core areas, with high levels of income and wealth, and peripheral regions, often specialized in low value-added sectors. One possible explanation for these patterns is that places are inherently different in terms of productivity; core urban areas are frequently characterized by the presence of natural advantages (i.e. first-nature advantages: large and deep harbors, or a central location in extensive plains with a highly productive agriculture system) that eventually generate agglomeration economies (i.e. second nature). However, the distribution of population is also subject to (non-economic) historical shocks that may have long-lasting effects (Rosenthal and Ross, 2015; Schumann, 2014). A possible consequence is that the distribution of population is *mislocated*, that is not optimally allocated across areas, and some high-(low)productivity sites end up being sub-optimally under-(over) populated. Spatial mislocation may also hamper the aggregate growth of a country (Hsieh and Moretti, 2015).

This paper analyzes the extent and the effects of the mislocation of population in Italy due to pirates' attacks launched from North-African shores against coastal locations (especially, in the south-west). Exposure to piracy and armed attacks started in the sixteenth century, when north-western Africa (today's Morocco, Algeria, Tunisia, and Libya) fell under the nominal influence of the Ottoman Empire, and lasted until the early nineteenth century, when France conquered modern-day Algeria. Owing to the fear of being attacked, Italian coastal localities in areas with a high probability of being raided lost their attractiveness for residents; the population grew relatively more in safer places (far from the coasts and difficult to assault because of their mountainous and rugged terrain). As a result, low productivity areas became relatively overpopulated. We proxy the probability of being attacked with the shortest sea distance between each Italian municipality and Tunis, which was the main port of departure for the raids. This is a reasonable proxy: it correlates well with actual slave-taking attacks that occurred in the period 1516-1798 for which we have data (Davis, 2003). We also provide a thorough robustness analysis for the findings by using alternative definitions of the security features of the havens and different "distance" measures for the likelihood of

2

being targeted by pirates. More importantly, our results are obtained by controlling for subsoil characteristics which take into account the productivity features of each location. As a placebo, we also show that the concentration of population in space was very different before the attacks, for instance during Roman times.

The effects of pirate attack on the distribution of Italy's population are shown to persist over time (Rosenthal and Ross, 2015). The magnitude of the impact measured for 1871 (the first year for which a complete census of Italian cities is available) is quite large even though the attacks had ceased more than 40 years before; the effect is still evident in 1951, though its magnitude diminishes by almost one-third. Thus, the effect persisted notwithstanding two world wars, and the exceptional wave of outward migration from the end of the nineteenth century to the 1920s. The impact ceases to be visible after 1971, erased by the massive south-to-north and rural-to-urban migration, a salient feature of the Italian industrialization process until the mid-1970s.

We then analyze the consequences of the spatial mislocation of the population for a number of local economic outcomes. Due to data constraints, all estimates refer to post-WWII censuses, when measured mislocation was quantitatively smaller than that registered 80 years before. Nonetheless, we find that the relative overpopulation driven by pirates' attacks in areas less suitable for economic activities correlates with a slower accumulation of human capital, higher migration flows mostly involving younger people, and overspecialization in subsistence agriculture. Finally, we present some suggestive calculations on the aggregate cost for Italy caused by the mislocation of population. The results show that the largest urban area in 1951 (Rome) would have been 15% larger (with 245,000 more inhabitants) in the counterfactual no-piracy scenario; urban primacy (the share of population living in a country's main city) would have been equal to 4.1% (rather than the actual 3.6%). We then use some recent estimates (Castells-Quintana, 2017) of the elasticity of GDP per capita growth to urban primacy to compute the loss due to the fact that piracy reduced the magnitude of agglomeration effects in Italy. We find that Italy has been converging on a lower steady-state GDP per capita level; the difference amounts to roughly 5% with respect to the no-piracy scenario.

Our paper relates to the literature on the consequences of historical shocks on city development and growth. Previous papers have exploited the fact that early access to trade routes shaped the systems of cities especially when urbanization was low; these systems were particularly persistent over time and survived even when the initial shock disappeared. Bleakley and Lin (2012) show that portage paths in the Appalachian region stimulated the concentration of people in their vicinity when they were significant trade routes. However, still today - when portage paths no longer provide location advantages – they are relatively overpopulated. Similar accounts are provided by Jedwab and Moradi (2016) and Jedwab et al. (2017), with reference to the colonial infrastructures in Africa, which became rapidly obsolete; Michaels and Rauch (2017) suggest that French cities were trapped in Roman locations with no coastal access. Differently from these papers, we do not exploit a natural or infrastructural original location advantage; instead, we refer to the consequences of predatory behaviors as in Nunn (2008) and Nunn and Puga (2012). Schumann (2014) also exploits an historical shock to illustrate population persistency. In particular, he shows how the relocation of German refugees after WWII determined a permanent increase of the inhabitants in those areas even after the ban on relocation was lifted. As Glaeser and Shapiro (2002) point out, the impact of warfare on urbanization processes may be either positive or negative. Military threats may generate a safe harbor effect and induce people to concentrate in urban areas, which are easier to defend; however, cities are also the main targets for military attacks and – in pre-industrial societies – their growth could be harmed by the destruction of transportation systems generally associated with wars. For instance, Dincecco and Onorato (2016) provide evidence in favor of the safe harbor effect by showing that military conflicts had an impact on the emergence of urban centers in Europe between 800 and 1799. Davis and Weinstein (2002) and Brakman et al. (2004) consider WWII and show that war-related damage had little impact on the distribution of population among cities, even in the short run. Similar findings are reported by Miguel and Roland (2011), who focus on the Vietnam War. Our paper complements the literature on military conflicts and the spatial distribution of population. While most studies on the effects of war-related damage concentrate on large-scale destruction that occurred in a very short period of time, we analyze low-intensity threats that lasted for an exceptionally long period.

The paper is structured as follows. The next section reviews the main historical accounts and highlights an example of two similar towns, differently exposed to the pirate threats. Section 3 describes our identification strategy and sets out the estimation results. It also reports robustness exercises intended to gauge the validity of our measures for the exposure to the attacks, and it discusses the role of competing explanations for our results. Section 4 looks at the persistence of the effect from the end of the attacks (1830) to the post-WWII period. Section 5 illustrates the consequences of being trapped in low-productivity places for the local economic development. Section 6 provides an attempt to analyze the impact of the attacks at aggregate level. Some implications of our evidence are discussed in the concluding section.

2. Historical background

This section provides a brief description of the pirate attacks. It then concretely illustrates the flight from the coast by considering the case of two similar towns differently exposed to the threat.

2.1 *Historical events*

Pirate activity started at the beginning of the sixteenth century as an attempt to avenge the Catholic conquest of Granada (Spain) in 1492. These initial attacks were mostly aimed against Spain, that, in self-defense, began to conquer the North-African coastal towns of Oran, Algiers, and Tunis. The Ottoman Empire – which acted as a protector for all Muslims at that time – reacted by helping local Berber populations to regain all North African cities under Spanish control from 1518 to 1529 (when Tunis fell). Ottoman sovereignty in those areas was only nominal (especially after 1659) as these territories were in fact anarchic military republics, which chose their own rulers and lived by plunder (Encyclopædia Britannica, 1911).

The entire Western Mediterranean coast was targeted by the raids, including cotemporary Spain, France, and Italy. However, as Davis (2003) points out, "[...] Italy was among the most thoroughly ravaged areas in the Mediterranean basin" due to Italian political fragmentation.[1] Ports of departure for raiders were Tunis, Tripoli, and

---

[1] After 1559, Italy was mostly under the influence of the Spanish crown, which had little incentive to provide protection for Italian shores.

Algiers. Figure 1 provides a historical representation of the barbary coast and particularly well illustrates the central position of Tunis as the port of departure and the overexposure of the south-west coast of the Italian peninsula, Sicily, and Sardinia. Konstam (2016) reports that, in 1620s, out of roughly 45 vessels devoted to raids in the entire Western Mediterranean, 34 were based in Tunis, 6 in Algiers and the remainder sailed from Tripoli.[2] The main targets in Italy were Sicily and the Tyrrhenian Italian coast, while the Adriatic coast was relatively sheltered by the presence of the Venetian navy.

[Figure 1]

At the end of the eighteenth century, piracy was still considered a major problem in the Western Mediterranean.[3] This called for military action: in several instances, British, Dutch, Sardinian, Neapolitan, and (even) US navies bombed Algiers, Tunis, or Tripoli during the so-called first Barbary War (1801-1805). Corsair activity based in Algiers did not entirely cease until France conquered the state in 1830; in the same period both modern-day Tunisia and Morocco fell under French influence, thus eliminating the attacks on Italian shores (Encyclopædia Britannica, 1911).

2.2 *A tale of two cities:* **Minturno** *and* **San Benedetto del Tronto**

The effects of pirate attacks on the distribution of population can be illustrated by comparing two towns. One is located in a heavily affected area (Minturno, Tyrrhenian coast) and another in a relatively sheltered location (San Benedetto del Tronto, Adriatic coast). Figure 2, panel (a) shows the position of the two towns on a map of Italy. As documented by Davis (2003), Minturno was heavily affected by piracy. Between the sixteenth and early eighteenth centuries, the town was raided several times; slaves were captured in nearby locations in 1588, 1644, and 1675. Pirate attacks on the area of San Benedetto del Tronto were instead comparatively milder. Raids were only recorded in the sixteenth century: by 1615 the town had expanded beyond its walls and toward the sea, indicating that the Adriatic was already safe in that period of time. Panels (b)

---

[2] At the same time, Algiers served as the main arsenal for pirates. In the 1620s the total number of vessels harbored in that city was slightly less than 60.

[3] For example, in 1798 the town of Carloforte (Sardinia) was raided and 900 inhabitants were enslaved and imprisoned for 5 years in Tunis (Paoletti, 2011).

and (c) of Figure 2 show the characteristics of the towns. Their physical geography is similar: both towns are by the sea and surrounded by hills; in both cases, the minimum altitude is zero and the maximum altitude is above 200 meters. From a geological point of view, the towns resemble each other in terms of average water content, depth to rock, and erodibility. Despite these similarities, the location of the two towns was markedly different, and it is still evident in current satellite images. The center of Minturno (the medieval town) is located on the top of a hill (141 meters) and distant from the shore (2.3 km), while the old town of San Benedetto del Tronto is much closer to the sea (700 meters) at a much lower altitude (4 meters). This comparison suggests that the threat of pirate attacks in the area of Minturno forced local inhabitants to choose an easier-to-defend location, even if this implied that areas more suitable in terms of economic development was left unsettled. In the next section, we develop this intuition by conducting a formal test to estimate mislocation.

[Figure 2]

3. Estimating mislocation

In this section we first discuss the empirical specification and the baseline results. We then present a number of robustness and placebo tests intended to probe the validity of our results also *vis-à-vis* some potentially competing explanations.

3.1. *Empirical specification*

The first set of data available on population at the municipality level refer to the period of the unification of Italy. At that time the attacks were over by less than half a century. We started our analysis by regressing municipality population in 1871 on a measure for the likelihood of being attacked, the characteristics of the territory that provided protection, and an interaction term between the two.[4] The estimating equation (where m stands for municipality) is the following:

---

[4] We also had 1861 data on a limited subsample of municipalities since the process of Italian unification was still incomplete. Results obtained using this subsample were similar to those referring to 1871 data (they are presented in Section 4).

(1) $\log(POP_{1871,m}) = \kappa + \alpha \Pr(attack_m) + \beta Protection_m + \gamma \Pr(attack_m) * Protection_m + \sigma X_m + \varepsilon_m$

Log population was regressed over a measure for the probability of being a target of pirate attacks, an index for the protection features of a given location, and an interaction term that captured the extent to which the concentration in sheltered localities was due to the fear of attacks, rather than other factors. γ represented our coefficient of interest.[5] We also added a number of town-specific characteristics intended to control for the determinants of local population different from the threat of assaults.

As regards the likelihood of being a target of a raid, Pr(Attack), we used the inverse distance between the municipality and the ports of departure; details on the construction of this measure are provided in Appendix A. Measuring the probability of receiving an attack with distance is consistent with Nunn (2008), who shows that, with reference to slave exports, deportations were decreasing in the distance to the final destination. We start by using the distance from Tunis, which is our baseline because the bulk of the attacks started from there (Section 2), and then provide robustness by using distance from alternative ports (Subsection 3.5). Our index for Protection takes several security features of the havens into account. Considering the characteristics of sixteenth-nineteenth century warfare, protection is higher for high-altitude locations. Moreover, ruggedness – sloping and irregular terrain – might provide additional defense against being raided (Nunn and Puga, 2012). In practice, we considered three variables: altitude of the municipality, difference between the highest and the lowest point in the municipal area, and terrain ruggedness.[6] We combined these three elements into a single measure derived with Principal Component Analysis (PCA; see Appendix A for details) that explained more than 80% of the overall variance of the three indicators. However, we also present evidence referring to the single components (Section 3.5). X_m include a set of city level controls. To proxy for the productivity characteristics of a site, we used the subsoil characteristics provided by the European Soil Database (ESDB). ESDB provides very detailed data (1-km-by-1-km) on a number of

---

[5] Antecedents for this specification can be found in the comparative advantages literature (Rajan and Zingales, 1998; Romalis, 2004; Nunn, 2007).
[6] The terrain ruggedness index expresses the amount of elevation difference between adjacent cells of a digital elevation grid: see Nunn and Puga (2012) for details.

geological characteristics. We aggregated this information at municipality level.[7] As in Combes et al. (2010), we considered 4 geological characteristics predicting the development of urban centers: Topsoil available water capacity (4 classes),[8] Subsoil available water capacity (5 classes), Depth to rock (5 classes), and Soil profile differentiation (4 classes). In addition, the location of population is severely affected by the frequency of earthquakes; indeed, high seismicity is likely both to destroy infrastructures (and buildings) and kill population. For this reason, we also used the seismic hazard of each municipality provided by the Italian Statistical Office.[9] Finally, we used a dummy equal to one if m was a capital city of a Sovereign State in the period 1789-1861 to control for possible overpopulation due to the presence of major administrative centers.[10]

There are two main identification assumptions for a consistent estimate of γ. The first is that the two characteristics are not correlated: this implies that the probability of being attacked is as good as randomly assigned with respect to the protection features of the site. Figure 3 shows that this condition was fulfilled in our data.

[Fig. 3]

The second assumption concerns the reduced-form flavor of equation (1). We hypothesize, in particular, that the inverse distance from Tunis is correlated with (i) the probability to be attacked, while (ii) it does not determine any other location determinant of population. As regards assumption (i) we show below (3.2) that the inverse distance from Tunis is positively correlated with the slave-taking attacks recorded in Italy between 1516 and 1798. Regarding assumption (ii) we provide a number of placebo exercises on Mediterranean ports (Subsection 3.3).

---

[7] Given that soil characteristics are usually discrete, we use the value that appears more often in each area.
[8] This feature proxies also for the presence of marshlands and swamps where malaria was particularly common in some areas of Sardinia, Tuscany, Latium, and Romagna. Until the 1950s, malaria was a major determinant of the location of population in Central and Southern Italy.
[9] Seismic hazard is a discrete measure ranging from 1 (high seismicity) to 4 (low seismicity).
[10] We consider the following capitals: Turin, Genoa, Milan, Venice, Lucca, Parma, Modena, Florence, Rome, Naples, and Palermo.

3.2 *Actual attacks and the distance from Tunis*

In order to validate our identification strategy, we used data on attacks by North-African pirates taken from Davis (2003), who focused on the major slave-taking events from 1516 to 1798 on the Italian coast: for 40 Italian cities the raids resulted in the enslavement of local people. As Davis pointed out, the data may suffer from severe underreporting because he collected data only on slave-taking for ransom.[11] As a result, attacks recorded in Davis's data were aimed against wealthy people, while much less is known about raids against poorer individuals for whom ransoms could not be demanded.[12] However, Davis's data are the only reliable source of information on pirate attacks and can give some interesting indications on the validity of our identification strategy. We aim to show that the inverse distance from Tunis – our proxy for the probability of being attacked – is correlated with actual pirate attacks. Our dependent variable ($Attack_m$) is a dummy equal to one if municipality m was raided according to Davis's data. We then estimate the following equation:

$$(2) \quad Attack_m = \kappa + \alpha \Pr(Attack_m) + \sigma X_m + \varepsilon_m$$

where $\Pr(Attack_m)$ is the inverse distance from Tunis, and X_m include the same set of controls of equation (1). Equation (2) is estimated with a linear probability model.[13] The results are presented in Table 1. The first column reports the α coefficient when no additional regressors are included; the second and the third columns show regression results when we include dummies for capital cities and quality of the soil. The estimation results validate the hypothesis that our measure for the probability of being attacked positively correlates with the actual attacks recorded by Davis (2003).[14]

[Tab. 1]

---

[11] This feature marks a stark difference with respect to slave-taking in Sub-Saharan Africa, where enslaved people were later employed in agriculture or mining; Nunn and Puga (2012).

[12] Davis (2003): "Impressive as these expeditions may have been – and they were the events most likely to go into the records – it is safe to assume that for each spectacular attack there were dozens, perhaps hundreds, of much smaller sorties […] against a handful of poor fishermen caught too far out at sea or a couple of village women snapped up while working in the fields."

[13] Using a probit or logit model delivers very similar results, with smaller standard errors.

[14] The standardized coefficient is between 0.026 and 0.036; the low magnitude might be due to the fact that recorded attacks are likely to be much fewer than those which actually occurred in that period.

3.3 *Baseline results and placebos*

The baseline results obtained from estimating equation (1) are set out in Table 2 (Table A1 provides the descriptive statistics). Column 1 shows the results when capital city and quality of soil dummies are excluded. The main effects have the expected signs; the greater the protection granted by (inhospitable) locations, the smaller the population. Towns with a high probability of being attacked, i.e. closer to Tunis, are instead somewhat larger, implying that southern and/or close-to-coast locations were able to sustain a larger population, probably because of better climatic and agricultural conditions. Our variable of interest has a positive sign and high statistical significance. The standardized coefficient is quite high: a s.d. increase in the interaction term is associated with a 47% rise in the s.d. of the municipal population. The magnitude of the interaction term diminishes to 29% once we include capital city and quality of soil dummies (Column 3) but the effect still remains substantial. The size of the local population can, in principle, depend also on the total land available for each municipality. In Column 4, we use 1871 population density as dependent variable. The results show that both the sign and the significance of the interaction variable are preserved.

To the extent that our evidence can be correctly attributed to pirate attacks – occurring from the sixteenth to the nineteenth century – we should fail to find an effect of our variable of interest on the spatial distribution of population *before the start of the raids*; indeed, during the Roman Empire, in which threats came from the North while Mediterranean shores were relatively safe, we should find that the estimate γ is negative. By using historical data from the Pleaides database of ancient places (Bagnall et al., 2016), we are able to select those that were Roman sites among the 7,687 Italian municipalities. 661 municipalities can be classified as previously Roman. As we did not have information on the respective populations, we used as outcome a dummy equal to 1 for the 1871 municipalities with Roman origins and re-estimate eq. (1) by means of a linear probability model. The results depicted in Column 5 show that the coefficient of interest enters with a negative, rather than positive, sign. This is not surprising since traffics with northern Africa under the Roman Empire were limited to commercial flows and no piracy systematically plagued the Mediterranean Sea; the negative sign can, indeed, be explained by the fact that at that time military threats consisted of Barbarian

invasions from the North, which made Italy's northern locations relatively more vulnerable and, consequently, less populated. During the period in which boats from northern Africa carried pirates, other ships sailed the Mediterranean. They were peaceful vessels, transporting goods and people for trade purposes (they were also targeted by the pirates). One of the most important ports of departure for peaceful ships was Marseille; Column 6 shows that if Marseille is considered as a (placebo) port of departure of pirate attacks, the interaction term is not statistically different from zero, thus indicating that our baseline results are due to military threats rather than standard trading routes.

[Tab. 2]

3.4. *Earlier (ninth-to-eleventh century) raids*

In the baseline estimates (referred to 1871) we have shown that – in areas characterized by a high probability of pirate raids – locations in internal and more rugged areas were overpopulated compared with areas with a low probability of being attacked. We have also shown that this was not the case in Roman times. Our interpretation of these results is that overpopulation was due to sixteenth-to-nineteenth century pirate attacks from Northern Africa. However, a major concern is the circumstance that pirate attacks against the Italian coast were also recorded between the eighth and eleventh centuries: the so-called "Arab Conquest" of the southern shores of the Mediterranean Sea. These attacks started after the occupation of Northern Africa in the seventh century; in particular, after the fall of Carthage (modern Tunisia) in 698. Sicily was gradually occupied (827-902) and raids were recorded (during the ninth and eleventh centuries) on the Tyrrhenian coast by ships leaving from Sicilian ports. Attacks were mostly for the purpose of pillaging villages and capturing people to be sold on slave markets, but other settlements were chosen as bases for further expansion (Gosse, 1933).[15] Arab raids came to a halt after the Christian Reconquista of Sicily (1061-1091) by the Normans.

---

[15] These settlements included southern France (La Garde - Freinet), Northern Campania (Traetto), and Bari.

In order to correctly interpret our baseline results, we have to rule out that overpopulated unproductive locations were already configuring the shape of the Italian urban system *before* the sixteenth century. Historical data on local populations are difficult to obtain. The best available dataset is the one created by Bairoch et al. (1988), which provides information on the population of European cities for all locations that – at least at one point in time between 800 and 1850 – had a minimum of 5,000 inhabitants.[16] This data source covers a very long time span. However, its reliability before 1300 is quite low (see: Bairoch et al., 1988). Therefore, our results that make use of the entire time span (800-1850) are reliable only from 1300 onwards. This implies that we can test the "Arab conquest" effects only 300 years after the end of attacks. In other words, we cannot check the contemporaneous effect of "Arab conquest" but we can assess whether – at the start of the new wave of attacks in the sixteenth century – the Italian urban system was still affected by the previous raids.

We created a dummy variable equal to one if the city was recorded in the Bairoch et al. (1988) data and used this dummy as dependent variable in equation (1).[17] Figure 4 shows the (standardized) estimated coefficient of the interaction term (coefficient γ of equation 1) for all years in which Bairoch et al.'s (1988) data are available.[18] For the period 800-1500 the estimated coefficients are very low and not statistically different from zero. The coefficient becomes positive and significant from 1600 onwards (pirate attacks started in the 1510s). This evidence confirms that at the beginning of the sixteenth century there was no impact of the "Arab conquest" and that mislocation of Italian population in 1871 was due to pirate attacks of the sixteenth to nineteenth centuries.

[Fig. 4]

3.5. *Additional robustness*

---

[16] Data are available for the following years: 800, 900, 1000, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800, and 1850.

[17] Population threshold in Bairoch et al.'s (1988) data could be very restrictive for our analysis. According to the 1871 Census, the average size of an Italian town was 3,550; the share of towns with more than 5,000 inhabitants was just 14%. It is for this reason that we do not use log(population) as dependent variable but as a dummy for larger locations; this approach is similar to the one adopted for Roman settlements in Table 2 (Column 5).

[18] All regressions are without location controls *X*.

As explained in Section 2, the majority of the boats set sail from Tunis. However, other ports in northern Africa were probably involved in pirate departures. This is the case of Algiers and Tripoli, according to Kontsam (2016). In Table 3, Columns 1 and 2 we show what happens to our results if we measure the probability of being attacked by using the inverse distance from Algiers and Tripoli instead of Tunis. As the two alternative maritime routes are highly correlated with our baseline path, this check should be taken *cum grano salis*. The results, in any case, are confirmed. We also calculated $\Pr(Attack_m)$ by using a weighted average of the probabilities of being attacked with respect to these three main ports of departure, with the weights suggested by the historical accounts of Kontsam (2016). The results are set out in Column 3.

Historical accounts suggest that adequate protective sites were provided by inland areas with high altitude and a rugged terrain. Our variable 'Protection' therefore considered these three elements jointly through a PCA routine. The first component, which we took as a proxy for Protection, was the only one with an eigenvalue greater than one and explained more than 80% of the common variance (see Table A2). As Italy has a very diversified territory, a pertinent question is whether the three elements should be jointly considered. Columns 4, 5, and 6 provide the regression results obtained for the baseline specification of Table 1, Column 3 when the individual components of our composite index for Protection are used. The coefficients on the interaction enter with very high significance. Note also that our results suggest that ruggedness and – to a lesser extent – altitude play the main role in driving the coefficient; this is in line with the findings of Nunn and Puga (2012).

[Tab. 3]

3.6 *Competing explanations*

An important issue is whether our findings might be explained by factors different from the fear of pirates. Some alternative explanations have already been considered through our set of covariates. For instance, subsoil characteristics should control for the agriculture-based explanations for urbanization, such as the adoption of the heavy plow (Andersen et al., 2016) or the start of potato cultivation (Nunn and Qian. 2011). Similarly, population dynamics triggered by earthquakes (Belloc et al., 2016) are

14

controlled for by our measure of local seismic hazard. Other competing explanations contrast with the ancillary evidence that we have provided. For instance, the fact that mislocation started to be observed in 1600 is at odds with the idea that our findings might be due to the (potentially asymmetric at the local level) effects of the Industrial Revolution in other European countries, which began in the 1760s.

Other explanations, however, might more subtly undermine our findings. We are particularly concerned that our proxy $\Pr(Attack_m)$, which has a clear north (low probability)/south (high probability) pattern, may also capture aspects related to differences among Italian regions that only by chance happen to be correlated with it. In 1871 – ten years after the unification of Italy – differences between southern and northern localities were probably even more pronounced than contemporary ones (Felice, 2013) because regions previously belonged to different nations. Therefore, post-unification population trends might confound our estimates. To tackle these additional concerns, our strategy was to use fixed effects at increasing levels of geographic details. We were thus able to compare – within progressively more homogenous areas – what happened to our coefficient of interest. This experiment, however, was not without costs. Since we worked on cross-sectional data, an important amount of variability in our variable of interest was differentiated away when introducing geographical fixed effects.[19] Table 4, Column 1 introduces a dummy for macroregions (5 dummies at NUTS1 level in the European Union Nomenclature). The results are very similar to those of the baseline. Column 2 adds regional (NUTS2 in the European Union nomenclature) dummies (we have 19 areas because Trentino-Alto Adige was not part of Italy in 1871). The results are only marginally modified. The inclusion of NUTS2 fixed effects also helps to consider the role of other explanations. For instance, Dincecco and Onorato (2016) suggest that for a long period of time (from the fall of Charlemagne's empire to the start of the Industrial Revolution; a period which largely overlaps with the one considered), cities developed because they were safe havens from military conflicts. To measure the strength of military threats, Dincecco and Onorato use 150 km x 150 km grid-scale cells, which roughly correspond to NUTS2 units; therefore, our estimates in Column 2 likely remove spatial population dynamics

---

[19] For instance, by adding NUTS2 fixed effect amount to differentiate away the NUTS2-average inverse distance to Tunis and the NUTS2-average protection features of the sites.

due to conflict exposure. In Column 3 we use instead dummies for pre-unification Italian states (we use 1859 borders). The shape of an urban system crucially depended on the size of each sovereign state (see: Cervellati et al. 2019); and military capabilities to deal with pirate attacks differed greatly among pre-unification states. Moreover, population mobility was generally much larger within, rather than among, sovereign countries, thus affecting the impact of military threats on the internal re-distribution of the population. Baseline results are confirmed (even in magnitude) by the introduction of pre-unitary states dummies. Finally, Column 4 controls for North-South (and East-West) differences by introducing a linear term for latitude and longitude. Also in this case, the results are unaffected.

[Tab. 4]

4. Long-run effects

Table 5 reports the estimation results by using as dependent variable municipal population measured at census dates from 1861 to 2001. The results are striking. The impact of pirate attacks is still detectable in 1971 (even though statistically significant up to 1961), almost one century and a half after the raids had terminated. The pattern of point estimates obtained is monotonically (and slowly) decreasing over time: for instance, in 1951 the effect estimated was roughly one-third of that measured in 1871. This persistence is noteworthy and consistent with the large body of literature on population dynamics across cities (see, Rosenthal and Ross, 2015, for a survey). During the period covered by Table 5, the spatial distribution of the Italian population was shocked by the two world wars. In addition, there was an exceptional wave of outward migration from the end of the nineteenth century to the 1920s. Finally, at the beginning of the twentieth century Italy experienced a first wave of industrialization (Castronovo, 1995) led by the regions of the north (with related rural-to-urban migration flows). On the other hand, the persistence might have been also due to the overall structural characteristics of Italy's economy from Unification to the aftermath of WWII, mainly agricultural with limited internal migrations, and the prolonged period of Fascist dictatorship, which banned migration (Andini et al, 2016). The effect estimated ceases to exist after 1971. This is not surprising as internal migrations, mainly south to north and rural to urban, were a distinctive feature of the Italian industrialization process

until the 1970s. The interaction term becomes negative and significant in 2001; as we show in the next section this is due to the fact that internal migration mostly affected younger individuals, thus leaving older people in the most affected areas. In the very long run, the consequential low birth rates are likely leave those municipalities *under*-populated.

[Tab. 5]

5. Effects on local economic development

Owing to the fear of attacks by pirates from Northern Africa, the population tended to concentrate in easy-to-protect locations in regions more exposed to raids. Unfortunately, highly defensible locations were also generally less suitable for economic activity. In general, inland locations had lower market access; high altitude reduced crop yields; sloping terrain was generally characterized by lower agriculture productivity. To check the extent to which pirate attacks caused local economic outcomes to deteriorate, we estimated the following equation:

(3) $Y_m = \kappa + \alpha \Pr(Attack_m) + \beta Protection_m + \gamma \Pr(Attack_m) * Protection_m + \sigma X_m + \varepsilon_m$

where Y_m is now an indicator denoting local economic characteristics. Data availability constrained our choices in terms of indicators and time span. We focused on four main outcome variables: (i) share of employees in agriculture, (ii) number of non-agricultural plants per capita, (iii) human capital, and (iv) ageing index. The first two refer the local economic structure in terms of industry specialization and entrepreneurship rate. The other two describe, instead, local population in terms of incentives to invest in education and age class structure. For the sake of simplicity, Tables 6 to 9 report the estimated coefficient (and its standardized version) for the interaction term only; we concentrate on three years: 1951 (when overpopulation was still sizable), 1971 (the first year in which overpopulation disappears), and 2001 (the last year in which comparable data are available; 1991 for plants). Table 6 presents the estimates when we used the share of agricultural employment as outcome variable. Estimation results show that in 1951 overpopulated municipalities were relatively more specialized in

agriculture; the magnitude is large, with a standardized coefficient of 0.35; this feature persisted in subsequent years. The fact that high-protection areas are also less suitable for high value-added agricultural production – due to altitude and ruggedness – implies that inhabitants of those areas mostly lived on subsistence farming. We subsequently analyzed (Table 7) the impact of relative overpopulation in areas with bad geography on plants per capita. Since this is only mildly negatively correlated with the share of agriculture, we interpret it as an indicator for the entrepreneurship rate of the local population.[20] In 1951 the interaction term was negative and significant   (-0.31 the standardized coefficient) and it remained roughly constant in subsequent years (-0.30 in 1991). [21]

[Tab. 6]

[Tab. 7]

Overspecialization in subsistence agriculture and low entrepreneurship rates in the "modern" sectors also reduced the incentives to accumulate human capital. Table 8 presents the results by using the share of individuals with at least a secondary school diploma, which is the only outcome variable that is consistently available from 1951 onwards.[22] Still in 1951 relatively overpopulated localities were characterized by a lower average level of schooling than other areas (-0.07 in standard deviation terms). The effect is still detectable (but smaller) in 2001, despite the great expansion of high school education in Italy as a consequence of the mid-1960s reforms of the compulsory school system . Outmigration from relatively overpopulated areas also changed the age

---

[20] Over the period 1951-1991, the correlation between share of agricultural employment and non-agricultural plants per capita was -0.19.

[21] Interestingly, interaction terms for both agricultural share and plants per capita remain negative and significant even when estimated relative overpopulation disappears after 1981. This result can be interpreted in light of a core-periphery model in a New Economic Geography framework. As explained by Ottaviano and Thisse (2004), industries characterized by imperfect competition, increasing returns to scale, and trade costs tend to concentrate in more densely populated areas (so-called 'Home Market Effect'); the depopulation pattern that we observe in previously relatively overpopulated areas (see Table 5) might have determined a rise in the comparative advantage of core areas in those industries. This implies that previously relatively overpopulated municipalities had to rely even more on the agricultural sector and registered a further decay in the number of (non-agricultural) plants per capita. This might explain why the impact survives after the end of the treatment.

[22] Data on tertiary education are available from 1971 onwards, with results much in line with those of Table 8 (results available upon request).

structure of the non-migrant residents. Table 9 presents the results by using the Ageing Index (i.e. the share of individuals aged over 65 years old) as outcome variable. Still in 1951, the age structure in affected municipalities was not particularly different from that of other towns. The migration flows that started in the 1950s, however, mostly involved younger individuals, thus determining a rise in the ageing index by almost 42% in terms of standard deviation.

[Tab. 8]

[Tab. 9]

6. Possible implications for aggregate effects

Relative overpopulation due to the fear of pirate attacks induced individuals to concentrate in areas less suitable for economic development. *Local* economic activity was hampered in terms of overspecialization in subsistence agriculture, lower entrepreneurship rate, slower accumulation of human capital, and, in the long run, permanent outmigration of younger population. In this section we take a broader approach and provide tentative evidence on the effects of mislocation on the Italian economy *as a whole*.

A large body of theoretical literature has shown that the internal distribution of economic activities impacts on the growth potential of a country (Baldwin et al., 2001; Fujita and Thisse, 2002; Baldwin et al., 2003; Accetturo, 2010; Fujishima, 2013). If knowledge spillovers in innovation are localized, the aggregate growth of a country may benefit from the spatial concentration of economic activities in core areas. Relative overpopulation in peripheral areas instead generates an incomplete agglomeration: a slower accumulation of knowledge in the early stages of development may imply that the country could conditionally converge to permanently lower steady state equilibrium. Empirical evidence confirms that urban agglomeration fosters growth, especially in the early stages of development (Henderson, 2003; Brülhart and Sbergami, 2009; Castells-Quintana, 2017). For the Italian case, piracy may have hampered the emergence of larger urban areas, with possible negative effects on steady-state GDP per capita levels.

To test this idea, we proceeded in two steps. First, we computed a counterfactual distribution of city-sizes in 1951 if piracy had not existed in the past; we calculated agglomeration indices on both actual and counterfactual distributions. Second, we used the elasticity of GDP per capita growth with respect to agglomeration (measured by urban primacy) computed by Castells-Quintana (2017) to assess the loss in terms of growth and steady-state GDP per capita level due to under-agglomeration. We used 1951 (rather than 1871) as a starting year for two main reasons. The first was that Castells-Quintana's estimates refer to the period 1960-2010; providing figures on the impact of under-agglomeration over the period 1871-2010 would be extremely evocative, but available elasticities of growth to agglomeration refer only to the post-WWII period and can hardly be extended to a longer time span. The second reason was that the agglomeration-growth nexus is strictly linked to industrialization and economic take-offs; before WWII, Italy was still an underdeveloped country, with a very large agricultural sector and a limited industrial base. This feature implies that the scope for growth-enhancing agglomerations was limited.

6.1. *Counterfactual city-size distribution*

To compute a counterfactual distribution, we hypothesized that the Italian total population was constant and we redistributed city-level populations assuming that all locations were equally affected by piracy.[23] Starting from equation (1), by definition, the actual log population of a city is equal to:

(4) $\log(POP_{1951m})^A = \hat{\kappa} + \hat{\alpha}\Pr(attack_m) + \hat{\beta}Protection_m + \hat{\gamma}\mathrm{I_m} + \hat{\sigma}X_m + \hat{\eta}GeoDummies_m + \hat{\varepsilon}_m$

where $I_m = \Pr(attack_m) * Protection_m$ of eq. (1); "Hats" define OLS estimations of Table 2, col. (3).

---

[23] The assumption that the population is constant could be quite strong because individuals in relatively overpopulated areas could choose to move to a different country rather than internally migrate. However, in the 1950s and 1960s Italian internal migrations were (for the first time) stronger than outmigrations (Ascoli, 1979).

Counterfactual population was computed by assuming that the term I_m is the same for all locations and equal to the sample average $\bar{I}$. In formula:

(5) $\log(POP_{1951m})^C = \log(POP_{1951m})^A - \hat{\gamma}[I_m - \bar{I}]$

A comparison between actual and counterfactual distribution is displayed in Table 10. It is apparent that the actual 1951 Italian system of cities was less agglomerated than the one which could have been observed without piracy; standard deviation in the distribution of log-population was lower; locations in the bottom percentiles were larger, while larger locations were comparatively smaller. For example, a city at the 95th percentile of the distribution was almost 7% smaller than the one in the counterfactual scenario. The Zipf coefficient is lower in the actual data, thus confirming that agglomeration was lower in 1951.[24] As an illustrative example, we have also reported the actual and counterfactual populations of Rome in 1951; according to our calculations, the largest Italian city could have been 15% more populous without piracy. We also report the difference between actual and counterfactual urban primacy (the share of the Italian population living in the biggest city) that approximates the relative size of the largest urban center in the country and is a commonly used indicator of urban agglomeration; according to our calculation, without piracy, urban primacy could have been 0.5 percentage points larger.

[Tab. 10]

6.2. *Implications for aggregate growth*

We can now compute the loss in terms of GDP per capita growth due to the lack of agglomeration. We use the elasticity of 5-years GDP per capita growth with respect to urban primacy provided by Castells-Quintana (2017) for the period 1960-2010; according to Castells-Quintana, for non sub-Saharan African countries, the coefficient is 0.0088, with a standard error of 0.003.[25] The loss for the Italian economy can be

---

[24] The Zipf coefficient is obtained by regressing the log rank of the city on the log population. The larger the coefficient, the more concentrated is the population. Standard errors are bootstrapped to take cross-sectional dependence into account.

[25] See Castells-Quintana (2017), table 2, col. (6). According to the author, the elasticity is even larger for European and Asian countries only.

calculated as 0.0088*0.6*10=0.05; between 1960 and 2010, the Italian lack of agglomeration determined a cumulated lower growth of 5 percentage points; Appendix B proposes a methodology to calculate the confidence interval for this estimate and shows that this figure is statistically significant. Owing to piracy, the Italian economy converges on a steady-state GDP per capita that is 5% lower than the counterfactual scenario.[26]

7. Conclusions

In this paper we have shown how historical shocks may have persistent effects on the spatial distribution of the population and, in turn, on the economic development of an area. We first presented evidence that, due to the fear of pirate attacks, the populations of some areas of Italy (especially south-west) concentrated relatively more in locations that were easier to defend but less productive. As a result, those areas recorded worse economic outcomes in terms of human capital and industry specialization; in the long run those towns were characterized by a marked out-migration. We also presented evidence that relative overpopulation in low productivity areas prevented the development of important urban centers with negative effects on aggregate incomes.

These results may have significant consequences in both positive and normative terms. From the positive perspective, we have shown that in the areas affected by pirate attacks the advantages of relative overpopulation in terms of agglomeration economies were largely exceeded by the disadvantages in terms of productivity; as a result, once the historical event that had determined concentration was over, those locations slowly depopulated. From the normative perspective, our results cast doubts on the economic foundations of the policies intended to halt population outflows from places characterized by low productivity and weak fundamentals.[27] Places experiencing de-population patterns might have been ones reverting to previous historically-driven

---

[26] Steady-state GDP per capita difference is computed as 0.5*(0.0088/0.0837), where 0.0837 is the coefficient of the lagged dependent variable (in log levels) of the Barro regression estimated by Castells-Quintana (2017).

[27] For example, the Italian government has recently proposed a program for the development of Internal Areas (Aree Interne), with the aim to resist depopulation and attract economic activities (see: http://www.dps.tesoro.it/aree_interne/ml.asp). Many of the areas eligible under this policy are exactly those that – because of the pirate threat – were overpopulated up to the mid 1970s.

overpopulation. Policies intended to keep people in bad locations could generate negative consequences on welfare at not only local but also aggregate level.

References

- Accetturo, A. (2010). "Agglomeration and growth: the effects of commuting costs." *Papers in Regional Science*, 89(1), 173-190.

- Amari, M. (1933). "Storia dei Musulmani di Sicilia. 2nd edn, modificata e accresciuta dall'autore, pubblicata con note a cura di CA Nallino. 3 vols." *Catania,* 1933.

- Andersen, T. B., Jensen, P. S., and Skovsgaard, C. V. (2016). "The heavy plow and the agricultural revolution in Medieval Europe." *Journal of Development Economics*, 118, 133-149.

- Andini, M., A. Dalmazzo, and G. de Blasio (2016). "The size of political jurisdictions: A model with some evidence from a fascist consolidation." Forthcoming *Journal of Comparative Economics*.

- Angrist, J. and J.-S. Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton university press.

- Ascoli U. (1979). *Movimenti migratori in Italia.* Bologna, Il Mulino.

- Bagnall, R. et al. (eds.) (2016). *Pleiades: A Gazetteer of Past Places.* <http://pleiades.stoa.org/> [Accessed: September 26, 2016 1:00 pm].

- Baldwin, R. E., P. Martin, and G. Ottaviano (2001). "Global income divergence, trade, and industrialization: The geography of growth take-offs." *Journal of Economic Growth* 6.1: 5-37.

- Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicoud (2003). *Economic geography and public policy*. Princeton University Press.

- Barro, R. J., and X. Sala-i-Martin (1991). "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics* 106.2: 407-443.

- Barro, R. J., and X. Sala-i-Martin (1992). "Convergence." *Journal of political Economy*: 223-251.

- Belloc, M., Drago, F., and Galbiati, R. (2016). "Earthquakes, religion, and

transition to self-government in Italian cities." *The Quarterly Journal of Economics*, 131(4), 1875-1926.

- Bleakley, H. and J. Lin. (2012). "Portage and path dependence." *The Quarterly Journal of Economics* 127.2: 587.

- Brakman S., H. Garretsen, and M. Schramm (2004). "The strategic bombing of German cities during WWII and its impact on city growth." *Journal of Economic Geography*, vol. 4, pp. 201-218.

- Brülhart, M. and F. Sbergami (2009). "Agglomeration and growth: Cross-country evidence." *Journal of Urban Economics* 65.1: 48-63.

- Castronovo, V. (1995). *Storia economica d'Italia: dall'Ottocento ai giorni nostri*. Vol. 619. Einaudi.

- Castells-Quintana D. (2017). "Malthus living in a slum: Urban concentration, infrastructure and economic growth." *Journal of Urban Economics* 98: 158-173.

- Cervellati, M., Lazzaroni, S., Prarolo, G., and Vanin, P. (2019). "Political Geography and Pre-Industrial Development: A Theory and Evidence for Europe 1000-1850" (No. 13719). CEPR Discussion Papers.

- Colley, L. (2004). *Captives: Britain, Empire, and the World, 1600–1850*, Anchor Books Edition, New York.

- Davis R. (2003). "Christian Slaves, Muslim Masters. White Slavery in the Mediterranean, the Barbary Coast, and Italy, 1500-1800." Palgrave-McMillan.

- Davis D and D. Weinstein (2002). "Bones, bombs, and break points: the geography of economic activity." *The American Economic Review*, vol. 92, pp. 1269-1289.

- Dincecco, M., and Onorato, M. G. (2016). "Military conflict and the rise of urban Europe." *Journal of Economic Growth*, 21(3), 259-282.

- Donaldson D. and R. Hornbeck (2016). "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics*, vol. 131, n. 2, pp. 799-858.

- Eeckhout, J. (2004). "Gibrat's law for (all) cities." *The American Economic Review* 94.5: 1429-1451.

- Encyclopædia Britannica (1911). *Barbary Pirates.*

- Faber, B. and C. Gaubert (2016). "Tourism and Economic Development: Evidence from Mexico's Coastline." No. w22300. National Bureau of Economic Research.

- Felice, E. (2013). *Perché il Sud è rimasto indietro.* Il Mulino.

- Fujishima, S. (2013). "Growth, agglomeration, and urban congestion." *Journal of Economic Dynamics and Control* 37.6: 1168-1181.

- Fujita, M. and J.-F. Thisse (2002). *Economics of agglomeration: cities, industrial location, and globalization*. Cambridge university press.

- Gardiner, B., R. Martin, and P. Tyler (2010). "Does spatial agglomeration increase national growth? some evidence from Europe." *Journal of Economic Geography*.

- Glaeser E. and J. Shapiro (2002). "Cities and Warfare: The Impact of Terrorism on Urban Form." *Journal of Urban Economics*.

- Gosse, P. (1933). *The history of piracy*. Tudor Publishing.

- Henderson J. V. (2003). "The urbanization process and economic growth: the so-what question." *Journal of Economic Growth*, n. 8, vol. 1, pp. 47-71.

- Hsieh, C. and E. Moretti (2015). *Why do cities matter? Local growth and aggregate growth*. No. w21154. National Bureau of Economic Research.

- Jedwab, R. and A. Moradi (2016). "The permanent effects of transportation revolutions in poor countries: evidence from Africa." *Review of economics and statistics* 98.2: 268-284.

- Jedwab, R., E. Kerby, and A. Moradi (2017). "History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya," forthcoming *The Economic Journal*.

- Konstam, A. (2016). *The Barbary Pirates 15th-17th Centuries*. Bloomsbury Publishing.

- Michaels, G., and Rauch, F. (2017). "Resetting the urban network: 117–2012." The *Economic Journal*, 128(608), 378-412.

- Miguel E. and G. Roland (2011). "The long-run impact of bombing Vietnam." *Journal of Development Economics*, vol. 96, pp. 1-15.

- Nunn, N. (2007). "Relationship-specificity, incomplete contracts, and the pattern of trade." *The Quarterly Journal of Economics*: 569-600.

- Nunn, N. (2008). "The Long-term Effects of Africa's Slave Trades." *The Quarterly Journal of Economics* 123.1: 139-176.

- Nunn, N. and D. Puga (2012). "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics* 94.1: 20-36.

- Nunn, N., and Qian, N. (2011). "The potato's contribution to population and urbanization: evidence from a historical experiment. " *The Quarterly Journal of Economics*, 126(2), 593-650.

- Onorato M. and M. Dincecco (2016). "Military conflict and the rise of Urban Europe." *Journal of Economic Growth*; 21(3), 259-282.

- Ottaviano, G. and J.-F. Thisse (2004). "Agglomeration and economic geography." *Handbook of regional and urban economics*, 4, 2563-2608.

- Paoletti, C. (2011). *La marina sabauda dal 1798 alla restaurazione*. Bollettino d'Archivio dell'Ufficio Storico della Marina Militare.

- Rajan, R. G. and L. Zingales (1998). "Financial Dependence and Growth." *American Economic Review*, LXXXVIII, 559-586.

- Romalis, J. (2004). "Factor proportions and the structure of commodity trade." *The American Economic Review* 94.1 (2004): 67-97.

- Rosenthal, S. and S. L. Ross. (2015). "Change and Persistence in the Economic status of Neighborhoods and Cities." *The Handbook of Regional and Urban Economics* 5.

- Schumann, A. (2014). "Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II." *American Economic Journal:*

Applied Economics 6.3: 189-205.

Figure 1. Barbary Coast

Sources: Antonio Zatta, "Le coste di Barbaria, ovvero i Regni di Marocco, di Fez, di Algeri, di Tunisi, di Tripoli", Altante Novissimo Vol.4, Venice 1794

Figure 2. Affected vs. Non-Affected towns: an example of two towns

(a) Location of the two towns

(b) Affected town: Minturno (the orange dotted circle plots the old town)

Altitude: 141mt
Altitude min: 0mt
Altitude max: 204mt
Distance to the coast: 2353mt
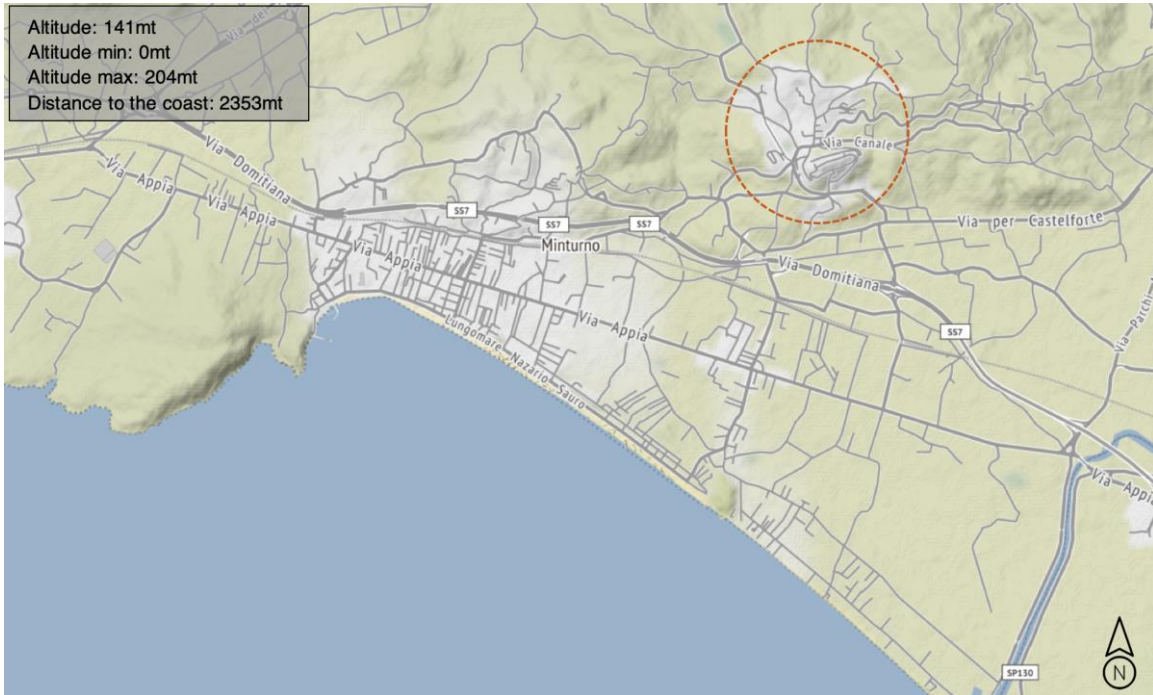
(c) Non-Affected town: San Benedetto del Tronto

Altitude: 4mt
Altitude min: 0mt
Altitude max: 283mt
Distance to the coast: 691mt

Figure 3. Correlation between Pr(Attack) and Protection
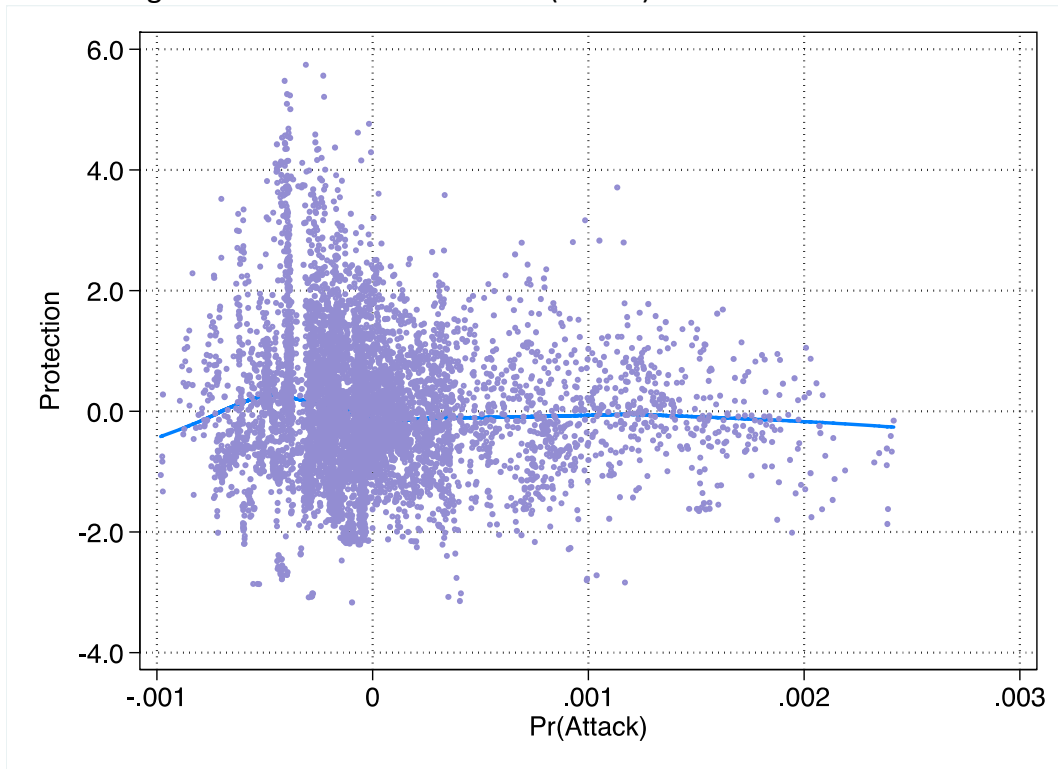
Notes: Authors' calculations on Istat data.

Each dot represents a municipality. The continuous line is the local polynomial estimate. *Pr(Attack):* inverse distance between the city and Tunis (see Appendix A for details). *Protection*: synthetic measure (first component) of Altitude, Slope of the ground, and Ruggedness; see table A2. All values in this graph represent the residuals of a regression of, respectively, *Protection* and *Pr(Attack)* on capital city and soil characteristics dummies.

Figure 4. The effects of earlier pirate attacks
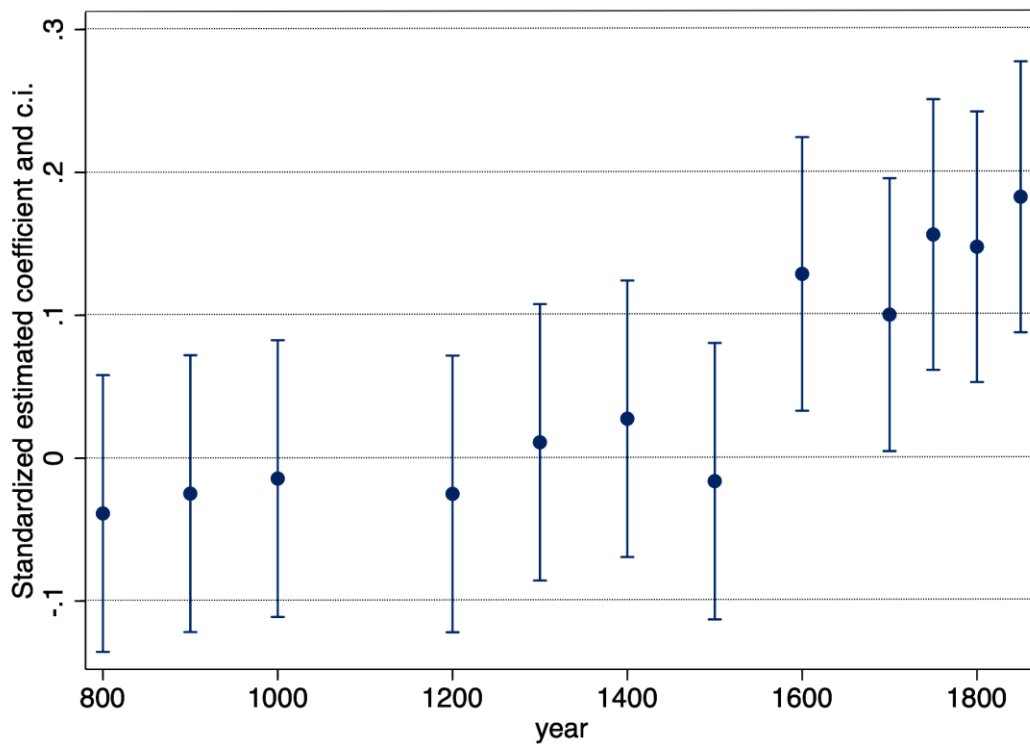
Notes: Authors' calculations on Istat and Bairoch et al (1988) data.

OLS regressions, see equation (1) and Subsection 3.4. Unit of observation: municipality. Number of observations: 7664. Dependent variable: dummy equal to one if municipality population is recorded in Bairoch et al. (1988) for each year. * (**) [***] denotes significance at the 10% (5%) [1%] level. The figure represents the estimated (standardized) coefficient for γ and its 99% confidence interval.

Table 1. Correlation between actual attacks and probability of attack

Dependent variable: Actual pirate attacks (1516-1798)	(1)	(2)	(3)
Pr(Attack)	3.4179**	3.3257**	4.6487**
	[1.4269]	[1.4397]	[2.0825]
Stand. coef.	0.0263**	0.0255**	0.0357**
Capital city dummy	NO	YES	YES
Quality of soil dummies	NO	NO	YES
R^2	0.001	0.003	0.007
No. Obs.	7664	7664	7664

Notes: OLS regressions, see equation (2). Unit of observation: municipality. Robust standard errors are in parentheses. * (**) [***] denotes significance at the 10% (5%) [1%] level. Dependent variable: dummy equal to one if the municipality was affected by pirate raids between 1516 and 1798 (see Davis, 2003). Pr(Attack) is taken to be the inverse distance from Tunis. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 2. The impact of pirate attacks on population, baseline estimates

Dependent variable:	Log population in 1871			Log population density in 1871	Placebo: Roman settlements	Placebo: distance from Marseille
	(1)	(2)	(3)	(4)	(5)	(6)
Protection*Pr(Attack)	223.5156***	228.6772***	137.0699***	89.0647***	-15.6916**	7.6586
	[19.0953]	[18.8763]	[18.0750]	[18.6262]	[6.7876]	[6.5859]
Stand. coef.	0.4651***	0.4759***	0.2852***	0.1988***	-0.1032**	0.0322
Protection	-0.3535***	-0.3567***	-0.2280***	-0.3357***	0.0077	-0.0827***
	[0.0226]	[0.0224]	[0.0224]	[0.0227]	[0.0077]	[0.0172]
Stand. coef.	-0.5898***	-0.5950***	-0.3803***	-0.6007***	0.0407	-0.1379***
Pr(Attack)	188.4432***	184.2283***	64.8490***	-403.8196***	39.4978***	-4.6523
	[21.3621]	[20.8767]	[21.8481]	[22.2419]	[8.0243]	[16.3051]
Stand. coef.	0.1134***	0.1109***	0.0390***	-0.2606***	0.0751***	-0.0047
Capital city dummy	NO	YES	YES	YES	YES	YES
Quality of soil dummies	NO	NO	YES	YES	YES	YES
R^2	0.051	0.086	0.204	0.294	0.065	0.197
No. Obs.	7664	7664	7664	7664	7664	7664

Notes: Authors' calculations on Istat, Pleaides and ESDB data.

OLS regressions, see equation (1). Unit of observation: municipality. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Roman settlement: dummy equal to one if the municipality hosted a Roman settlement as recorded in the Pleaides database. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis; in column (6) Pr(Attack) is the inverse distance from Marseille (see Appendix A for details). Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 3. The impact of pirate attacks on population, robustness on Pr(Attack) and Protection

Dependent variable: Log population in 1871	(1) Algiers	(2) Tripoli	(3) Weigh. Avg.	(4) Altitude	(5) Slope	(6) Ruggedness
Protection*Pr(Attack)	444.8953***	113.8084***	164.4837***	0.6821***	0.3524***	1.1653**
	[52.5680]	[30.9927]	[21.5078]	[0.0724]	[0.0385]	[0.4891]
Stand. coef.	0.8552***	0.2188***	0.3162***	0.0592***	0.0306***	0.1013**
Protection	-0.4710***	-0.1518***	-0.2473***	-0.0015***	-0.0003***	-0.0039***
	[0.0476]	[0.0260]	[0.0246]	[0.0001]	[0.0001]	[0.0006]
Stand. coef.	-0.7857***	-0.2533***	-0.4126***	-0.0991***	-0.0225***	-0.2635***
Pr(Attack)	-20.5890	256.9424***	78.1857***	-171.7886***	-169.0011***	-40.3088
	[64.7592]	[44.9272]	[26.1345]	[35.0734]	[36.4406]	[39.9772]
Stand. coef.	-0.0103	0.1290***	0.0393***	-0.1032***	-0.1015***	-0.0243
Capital city & Quality of soil dummies	YES	YES	YES	YES	YES	YES
R^2	0.205	0.203	0.204	0.221	0.204	0.208
No. Obs.	7664	7664	7664	7664	7664	7664

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (1). Unit of observation: municipality. Robust standard errors are in parenthesis. * (**) [***] denote significance at the 10% (5%) [1%] level. Pr(Attack) is taken to be the inverse distance from the localities in column (see Appendix A for details) in columns (1) to (3) and Tunis for columns (4) to (6). Protection is the first PC, as described in Table A2 in columns (1) to (3) and the single components in column for columns (4) to (6). Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 4. The impact of pirate attacks on population, geographical controls

Dependent variable: Log population in 1871	(1)	(2)	(3)	(4)
Protection*Pr(Attack)	136.1001***	102.7317***	122.3244***	123.2972***
	[18.4748]	[18.2405]	[18.3790]	[18.3313]
Stand. coef.	0.2832***	0.2138***	0.2546***	0.2566***
Protection	-0.2162***	-0.1523***	-0.1984***	-0.2029***
	[0.0228]	[0.0229]	[0.0228]	[0.0230]
Stand. coef.	-0.3608***	-0.2541***	-0.3310***	-0.3385***
Pr(Attack)	25.7430	350.1488***	12.4703	-11.9060
	[31.6937]	[66.2448]	[23.4367]	[54.1716]
Stand. coef.	0.0155	0.2107***	0.0075	-0.0072
Geographical controls	NUTS1	NUTS2	1859 Italian States	Latitude and Longitude
Capital city & Quality of soil dummies	YES	YES	YES	YES
R^2	0.212	0.279	0.237	0.207
No. Obs.	7664	7664	7664	7664

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (1). Unit of observation: municipality. Robust standard errors are in parentheses. * (**) [***] denotes significance at the 10% (5%) [1%] level. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A for details). Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard. Geographical controls: "NUTS1" for NUTS1 (ripartizioni) areas (5 dummies); "NUTS2" for NUTS2 (regioni) areas (19 dummies); "1859 Italian States" for pre-unitary (1859) Italian States borders. "Latitude and Longitude" for latitude and longitude included in linear way.

Table 5. The impact of pirate attacks on population, various Census dates

Dependent variable: Log population in…	Protection*Pr(Attack)	Standard error	Stand. coef.
1861	132.0701***	[18.8776]	0.2738***
1871	137.0699***	[18.0750]	0.2852***
1881	135.4172***	[18.3028]	0.2800***
1901	131.5714***	[18.8754]	0.2654***
1911	123.1528***	[19.0786]	0.2443***
1921	116.9596***	[19.4893]	0.2261***
1931	102.7895***	[18.3595]	0.1929***
1936	101.1051***	[18.2456]	0.1863***
1951	63.4142***	[18.6264]	0.1121***
1961	35.9349*	[19.3437]	0.0604*
1971	9.3502	[20.7254]	0.0145
1981	-15.5263	[21.8323]	-0.0230
1991	-35.5099	[22.3770]	-0.0510
2001	-47.0130**	[23.0942]	-0.0661**

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (1). Unit of observation: municipality. Number of observations: 6642 for the year 1861 and 7664 for all remaining years. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A for details). All regressions include capital city and quality of soil dummies. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 6. Share of Agricultural employment

Dependent variable: Share of agricultural empl.	Protection*Pr(Attack)	Standard error	Standardized coef.
1951	4626.8441***	[435.0971]	0.3513***
1971	3853.3138***	[358.4113]	0.3525***
2001	1682.8555***	[206.0643]	0.3369***

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (2). Unit of observation: municipality. Number of observations: 7414 for 1951, 7615 for 1971, 7664 for 2001. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Share of agriculture: share of employees in agricultural activities. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A, for details). All regressions include capital city and quality of soil dummies. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 7. Plants per capita

Dependent variable: Plants per capita	Protection*Pr(Attack)	Standard error	Standardized coef.
1951	-10.0385***	[1.0773]	-0.3148***
1971	-1.2204***	[0.3716]	-0.1286***
1991	-4.2148***	[0.5145]	-0.2975***

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (2). Unit of observation: municipality. Number of observations: 7660 for all years. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Non agricultural plants per capita: ratio between non agricultural plants and total population. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A, for details). All regressions include capital city and quality of soil dummies. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861.Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 8. Human Capital

Dependent variable: Share of individuals with at least a secondary school diploma degree	Protection*Pr(Attack)	Standard error	Standardized coef.
1951	-51.1850**	[25.0208]	-0.0744**
1971	-71.3666	[47.0338]	-0.0572
2001	-233.9050*	[127.4960]	-0.0675*

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (2). Unit of observation: municipality. Number of observations: 7414 for 1951, 7615 for 1971, 7664 for 2001. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Human capital: share of individuals with at least a secondary school diploma. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A, for details). All regressions include capital city and quality of soil dummies. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 9. Ageing Index

Dependent variable: Ageing rate	Protection*Pr(Attack)	Standard error	Standardized coef.
1951	120.0061*	[69.3565]	0.0531*
1971	1496.7733***	[134.6237]	0.3246***
2001	3175.4646***	[237.0363]	0.4218***

Notes: Authors' calculations on Istat and ESDB data.

OLS regressions, see equation (2). Unit of observation: municipality. Number of observations: 7414 for 1951, 7615 for 1971, 7664 for 2001. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. Ageing index: share of individuals with more than 65 years old over total population. Protection is the first PC, as described in Table A2. Pr(Attack) is taken to be the inverse distance from Tunis (see Appendix A, for details). All regressions include capital city and quality of soil dummies. Capital city dummy: dummy equal to one if the municipality was a the capital of a sovereign state in the period 1789-1861. Quality of soil dummies include: Topsoil Available Water Capacity, Subsoil Available Water Capacity, Depth to Rock, Soil Profile Differentiation, and Seismic Hazard.

Table 10. Actual and counterfactual city-size distribution in 1951

	Actual (A)	Counterfactual (B)	Difference (A)-(B)
St. dev. log pop.	1.040	1.077	-0.037[a]
p5 log pop.	6.303	6.221	0.081**
p20 log pop.	7.087	7.047	0.039**
p50 log pop.	7.895	7900	-0.005
p80 log pop.	8.713	8.739	-0.025
p95 log pop.	9.537	9.604	-0.066**
Zipf coefficient	-0.900	-0.867	-0.033[b]
(bootstrapped s.e.)	*(0.008)*	*(0.007)*	*(0.013)*
Pop. of Rome	1,651,394	1,896,267	-244,873[c]
Urban primacy	3.561	4.089	-0.528[c]

Notes: Authors' calculations on Istat data.

Unit of observation: municipality. Number of observations: 7664. Counterfactual scenarios are computed according to equations (3) and (4). [a] F-test for differences in standard deviations significant at 5%. * (**) [***] denotes significance at the 10% (5%) [1%] level for a quantile regression (bootstrapped standard standard errors). [b] The coefficient, significant at 5%, is obtained by pooling actual and counterfactual log population sizes and regressing a fully interacted Zipf equation. [c] Significance levels not provided.

Appendix A: Data construction, Descriptive statistics, and Principal Component Analysis

The probability of being attacked was measured by the inverse distance from Tunis. To measure this distance, we took two main issues into account:

- Transportation by sea was definitely faster than any other means of transport il the introduction of railways or internal combustion engines (Donaldson and Hornbeck, 2016);
- The orientation of the Italian peninsula is NW-SE, while Tunis is located SW.

In order to approximate the probability of being attacked by pirates from Tunis, we proceeded as follows:

1. we calculated the distance between each location c and the closest municipality located by the sea cS. This distance was denoted by D1. A could be equal to zero if c was located by the sea (i.e. c and cS coincided).

2. we calculated the minimum distance between cS and Tunis; this distance is denoted by D2. For this purpose, we considered three groups of localities on the Italian coast:

   a. Group A: these localities were characterized by the feature that a straight line between them and Tunis did not intersect with either the mainland or any other major island (Sicily and Sardinia).

   b. Group B: all other localities for which straight lines to Tunis traverse a major landmass. For this group we considered the following turning points to bypass major landmasses:

      i. For cS localities on the Adriatic coast: Otranto, Capo Passero (i.e. D2 was calculated as follows dist(cS, Otranto)+dist(Otranto, Capo Passero)+dist(CapoPassero, Tunis)).

      ii. For cS localities on the eastern Ionian coast: Capo Passero (i.e. dist(cS, Capo Passero)+dist(CapoPassero, Tunis)).

      iii. For cS localities on the western Ionian coast: Capocolonna, Capo Passero (i.e. dist(cS, Capocolonna)+dist(Capocolonna, Capo Passero)+dist(CapoPassero, Tunis)).

      iv. For cS localities on the northern Sicilian coast and the southern Tyrrhenian coast: San Vito Lo Capo (i.e. dist(cS, San Vito Lo Capo)+dist(San Vito Lo Capo, Tunis)).

v. For *cS* localities on the northern Sardinian coast: La Maddalena
(i.e. dist(*cS*, La Maddalena)+dist(La Maddalena, Tunis)).

3. The probability of being attacked was equal to 1/(D1+D2).

The spatial distribution of turning points is shown in Figure A1.

Figure A1. Turning points

The same procedure was used (with different turning points) when considering the distances to Algiers, Tripoli, Barcelona, and Marseille (Table 2). Figure A2 shows the distribution of the distances between each Italian municipality and Tunis or Marseille.

Figure A2. Distances to Tunis or Marseille

(a) Tunis   (b) Marseille
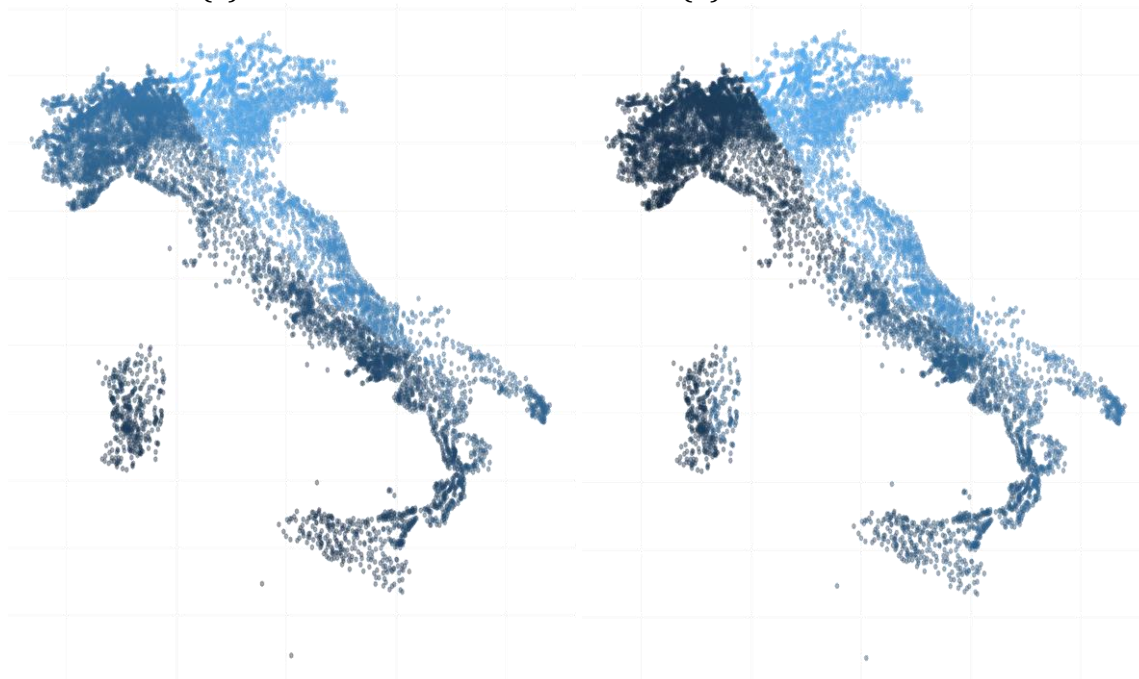
Notes: Authors' calculations on Istat data.

Each dot represents a municipality. Darker colors correspond to shorter distances.

Table A1. Descriptive statistics

	Mean	Standard Deviation
Protection:		
Altitude (continuous, in meters)	342.273	282.370
Slope (continuous, in meters)	617.392	619.445
Terrain ruggedness index	61.939	60.343
Capital city	0.002	0.040
Inverse distance from Tunis (continuous, 1/km)	0.001	0.001
Roman settlement (dummy 0-1)	0.086	0.280
Log(population-1871) (continuous)	7.658	0.885
Log(population-1951) (continuous)	7.933	1.041

Notes: Authors' calculations on Istat and Pleaides data.

Table A2. Principal Components results

	PC1	PC2	PC3
Scoring coefficients:			
Altitude	0.526	0.850	0.015
Slope	0.602	-0.360	-0.712
Terrain ruggedness index	0.600	-0.384	0.702
Eigenvalues	2.446	0.446	0.107
Explained variance	0.816	0.149	0.036

Notes: Authors' calculations on Istat and Pleaides data.

Appendix B: Confidence intervals for the estimation of aggregate effects

In what follows, we describe a methodology with which to compute confidence intervals for the estimation of aggregate effects.

As we saw in Section 6, the aggregate loss in terms of GDP per capita is the product between (a) the Castells-Quintana (2017) elasticity of GDP per capita growth with respect to urban primacy and (b) the difference in terms of urban primacy between the actual and the counterfactual scenario.

Regarding (a), confidence intervals are easy to obtain by using the standard errors provided by the paper. The 95% confidence interval for a point estimate of 0.0088 with a standard error of 0.003 is **a'**=[0.003; 0.015].

Confidence intervals for the component (b) are instead less trivial:

- We take the confidence intervals for the quintile regression at the 95th percentile between actual and counterfactual distributions of table 9. The point estimate at the 95th percentile is 0.066 (in log points), with a 95% confidence interval of [0.005; 0.141]. By using anti-log operator, this implies that, in percentage points, the point estimate is 0.068 and the confidence interval is **c'**=[0.005; 0.151].

- We calculated a proportionality rule between point estimate and the extremes of the confidence intervals as **d'**=**c'**/0.068**=**[0.005/0.066; 0.151/0.066]=[0.076; 2.288].

- We multiplied vector **d** by the difference in population for Rome in 1951 between actual and counterfactual scenario (244,873 inhabitants, see table 10). We obtained a vector **e'**=[18,610; 560,240] which represented the confidence interval for the difference in total population between the counterfactual and the actual scenarios; we could not reject the hypothesis that the population of Rome may have been 18,610 or 560,240 inhabitants larger.

- We divided vector **e** by the total population in Italy in 1951 (46,445,764) to obtain a confidence interval for the increase in urban primacy when we compared the counterfactual and the actual scenarios: **b'**=[0.040, 1.206]. This implies that, in the no-piracy scenario, we cannot reject an increase in urban primacy between 0.040 and 1.206 percentage points.

- Confidence intervals for the aggregate effect over the period 1960-2010 (ten 5-years periods) can be calculated as 10***ab'**:

$$10 * ab' = \begin{bmatrix} 0.0012 & 0.0362 \\ 0.0060 & 0.1809 \end{bmatrix}$$

Confidence intervals show that we cannot reject the hypothesis that cumulated growth rate in GDP per capita could have been 0.001 to 18 percentage points larger over the period 1960-2010. The fact that all extremes are positive implies that the point estimate of 0.5 percentage points is statistically significant.