STUCK IN THE WRONG PLACE: LONG-TERM CONSEQUENCES OF PIRATE ATTACKS IN ITALY[•]

Antonio Accetturo, Michele Cascarano, Guido de Blasio
Bank of Italy

Dic. 2, 2016

First and very preliminary draft, not for circulation

Abstract
From the VIII to early XIX century, coastal areas in Italy (especially, the South) were raided by pirates coming from the shores of Northern Africa. This paper documents that to protect themselves, residents of coastal places relocate inwards to mountainous and rugged areas. It also shows that the relocation had persistent effects. Overpopulation in areas less suitable for economic development determined that affected areas remained underdeveloped for a long period after the pirates' threat was over. We also present suggestive evidence that mislocation of population had negative aggregate effects on the economy.
JEL Classification: R1, N9, O1
Keywords: City structure; Historical shocks; Local development

[•] We wish to thank Vernon Henderson and the participants to a Bank of Italy workshop (April, 2016) for early discussions on the topic. The views and the opinions expressed in this paper are those of the authors and do not necessarily reflect those of the institution they are affiliated.

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 of the possible explanations for these patterns is the fact that cities are inherently different in terms of productivity; core urban are frequently characterized by the presence of some natural advantages (i.e. first nature advantages: large and deep harbors or the location at the center of large plains with a highly productive agriculture) that eventually evolved in agglomeration economies (i.e. second nature). In this vein, Bleakley and Lin (2012) provide a suggestive explanation of the evolution of urban centers in the Appalachian region in the US.

However, the distribution of population is also subject to (non-economic) historical shocks that can have long lasting effects (Rosenthal and Ross, 2015; Schumann, 2014). A possible consequence is that the distribution of population is *mislocated*, i.e. not optimally allocated across areas, and some high-(low)productivity sites end up to be suboptimally under-(over) populated. In this vein, Michaels and Rauch (2016) provide a nice comparison between the English and French urban systems. Spatial mislocation might also have negative effects on the aggregate growth of a country (Hsieh and Moretti, 2015).

The aim of this paper is to analyze the causes and the effects of the mislocation of population in Italy. We consider, in particular, the role of the attacks from the north-African pirates that coastal places (especially in the Southern-West coast) experienced in Italy over two different waves from and for very long period of time (i.e. from VIII century with the fall of the Byzantine Empire as a naval power to early XIX century when northern Africa, Tunis in particular, fell under the French influence). Due to the fear of being attacked, coastal places lost their attractiveness for residents who moved towards inner locations, far from the costs and difficult to assault because positioned on mountainous and rugged territories. As a result, these low productivity territories ended up in beign relatively overpopulated. These findings are robust to alternative definitions of the security features of the havens and using different measures for the likelihood of being targeted by pirates. More importantly, our results are obtained by using region fixed effects, which differentiate out all the features that permanently differ across Italian regions and that might alternatively explain inner and awkward settlements. We also show that the concentration of population over the space was very different before the VIII century: Roman cities were not located in places that ensured protection from the pirates, as there were no raids before the VIII century.

The effects of pirate's attack on the distribution of Italy's population are shown to persist overtime. The impact measured for 1871 (the first year for which a complete census of Italian cities is available) is still evident, tough its magnitude is reduced by a half, in the 1951 distribution of population. The effect persisted notwithstanding two world wars, the exceptional wave of outward migration from the end of the XIX century and the twenties of the XX century. The impact ceases to exist after 1981, wiped away by the massive south-to-north and rural-to-urban migration, which went hand in hands with the Italian industrialization process up to middle 1970s.

We then analyze the consequences of spatial mislocation of population induced by the pirate attacks on a number of economic outcomes. Due to the data availability constraint, all estimates refer to post-WWII censuses when measured mislocation was quantitatively smaller than the one registered 80 years before. This notwithstanding, we find that overpopulation derived by pirates' attacks in areas less suitable for economic activities determined a worse labor market outcomes, slower accumulation of human capital, and an over specialization in subsistence agriculture; out-migration in the period 1951-1981 also determined a sizable and persistent increase in the aging index with possible long-run consequences on the future development of those areas.

The reduced-form nature of our estimates does not allow to fully capture the aggregate consequences of such mislocation on aggregate economic growth.[1] However, we provide some suggestive evidence of the effects of our historical shock on the city size distribution and income levels. We show that highly mislocated provinces (the equivalent of a county for the US) did not observe the emergence of major urban

[1] To have an estimate of an aggregate effect we should resort to mixed reduced form-structural models like Kline and Moretti (2014) or Faber and Gaubert (2016).

centers; such feature prevented the full development of agglomeration economies and determined lower levels of average value added per capita at provincial level.

This paper relates to the literature on the consequences of historical shocks on city development and growth. Bleakley and Lin (2012) show that portage paths in the Appalachian region stimulated the concentration of people nearby when they were relevant trade routes. However, they remain today - when portage paths do not provide location advantages any longer - places relatively overpopulated. Similar stories are provided by Jedwab and Morandi (2014), with reference to the colonial railroads in Africa, which became quickly obsolete, and Michaels and Rauch (2016), who suggest that French cities were trapped in Roman locations with no costal access. Schumann (2014) shows how the relocation of German refugees after WWII determined a permanent increase in population in those areas even after the ban for relocation was lifted. Differently from the previous literature on path-persistence of population settlements, we do not exploit a natural or infrastructural original location advantage, but we refer to the consequences of predatory behaviors like in Nunn (2008) and Nunn and Puga (2012); we also provide evidence of long-run effects not only on population but also on other economic variables.

The paper is structured as follows. Next Paragraph provides the reader with the main historical accounts. Paragraph 3 introduces our baseline estimations and illustrates a number of robustness exercises. Paragraph 4 looks at the persistence of the effect from the end of the attacks to the period post WWII. Paragraph 5 illustrates the consequences of being trapped in inefficient places for the economic fortunes of Italy's territories. Some implications of our evidence are discussed in the concluding paragraph.

2. Historical background

From VIII to early XIX centuries Italy’s coasts were subject to pirate attacks coming from the shores of Northern Africa. There were two main waves. The first relates to the period VIII-XI centuries; the second started in the XVI century and finished in the early decades of the XIX century. These two waves were inherently different in terms of intensity and purposes.

The first was carried out during the great expansion of Arab domination in Africa and, in part, in Europe. The aim of these attacks in Italy was to create strategic outposts, which may serve as bases for a subsequent conquest by Arab Caliphs (Amari, 1933). Regular raids against Italian shores started after the fall of Carthage (modern Tunisia) under Arab conquerors in 698 AD. Between 703 and 752 Sicily was raided nine times by Arab fleets. An organized invasion of the island started in 827 and was finally accomplished 75 years later. While Sicily was occupied and colonized, in IX and XI centuries raids were registered on the entire Tyrrhenian coast (west coast of Italy) from ships leaving from Sicilian ports. Attacks were mostly aimed at pillaging villages and capturing people to be sold on slave markets but other settlements were chosen as bases for further expansion (Gosse, 1933).[2] Some raids went deep in the interior parts of the peninsula; for example Arab predators were registered in inner Latium and Umbria along with some areas of western Alps. Attacks on the Adriatic coast were rarer and mostly concentrated in the Southern parts: the most Northern point of an Arab attack on the Adriatic shore was Ancona (848 AD). Arab raids finished after the Christian *Reconquista* of Sicily (1061-1091) by the Normans.

The second wave started at the end of the XVI century; it was not aimed at occupying or colonizing territories but just predate kidnap people for ransom (Colley, 2004). This phase started on 1587 when the Barbary coast (i.e. modern Western Libya, Tunisia, Algeria, and Morocco) fell under the nominal Ottoman sovereignty but was actually governed by local rulers which chose to live by plunder (Encyclopædia Britannica, 1911). Starting ports for raiders were generally Tripoli, Tunis, and Algiers; 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 remaining sailed from Tripoli.[3] Even in these cases 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. At the end of the XVIII century, Barbary piracy was still considered a

[2] These settlements included southern France (La Garde - Freinet), Northern Campania (Traetto), and Bari.
[3] At the same time Algiers mostly served as the main arsenal for Barbary navy. In 1620s the total number of vessels harbored in this city skimmed 60.

major problem in Western Mediterranean.[4] This called for several military interventions: in several instances British, Dutch, Sardinian, Neapolitan, and (even) US Fleets bombed Algiers, Tunis, or Tripoli. 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 the French influence, thus greatly reducing the attacks on Italian shores (Encyclopædia Britannica, 1911).

3. Estimating mislocation

The first set of data available on population at the municipality level refers to the period of the unification of Italy.[5] At that time the attacks were over by half a century. We start by regressing municipality population in 1871 on a measure for the likelihood for being attacked, the characteristics of the territory that provide protection and an interaction term between the two.[6] The estimating equation (where m stands for municipality) is the following:

$$(1)\ \text{Log}(Pop\_1871_m) = k + \alpha\ Pr(attack)_m + \beta\ Protection_m + \gamma\ (Pr(attack)_m{*}\ Protection_m) + \text{Geo dummies}_m + \varepsilon_m$$

As for the likelihood of being a target of a raid, we use the inverse geodetic distance between the municipality and the shores of departures. Measuring the probability of receiving an attack with distance is consistent with Nunn (2002) that shows that, with reference to the slave exports, deportations were decreasing in the distance to the final destination. We start by using the distance from Tunis, which is our baseline as the bulk of the attacks departed from there (see Section 2) and then provide robustness by using

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

[5] Data constraints therefore do not allow us to distinguish between the two waves of the pirate attacks. Our estimates have to be interpreted are reflecting both waves. To the extent that the relocations occurring during the first wave (that ended at the end of the XI century) were completely reversed by the end of the XVI century (were raids started aging with the second wave), our estimates will be capturing only the impact of the second wave only. We believe that it is unlikely, however, that there was a complete reversal, given the results on persistency that we document in the paper.

[6] We also have 1861 data on a limited subsample since the process of Italian unification was still incomplete. Results obtained using this subsample are similar to those referring 1871 data (see Section 4).

both alternative and placebo measures of being attacked (see Section 3.1). Our index for Protection takes into account several security features of the havens. For attacks coming from the seaside protection is higher for internal and high-altitude locations. Moreover, ruggedness – sloped and irregular terrains – might provide additional defense from being raided (see: Nunn and Puga, 2011). The three elements are combined into a single measure derived with Principal Component. However, we also present evidence referring the single components (see Section 3.2).

Parameters α and β (main effects) capture, respectively, the influence of geographic location (like, for example, market access) and orographic characteristics on the size of the city.

Geo dummiesm include a set of geographical fixed effects. These are aimed at controlling for possible omitted variables (e.g. common economic conditions or institutions) that might influence the location of individuals in certain areas.[7]

Our coefficient of interest is γ, the parameter for the interaction term. It captures the extent to which the concentration in sheltered localities is due to the fear of attacks, rather than by other factors.

In this equation overpopulation is explained by the interaction between a location characteristics (inverse distance from Tunis) and an orographic characteristics of the city. Antecedents for this specifications can be found in the comparative advantages literature (Rajan and Zingales, 1998; Romalis, 2004; Nunn, 2007; Accetturo et al., 2015). The identification assumption for a consistent estimate of γ is that the two characteristics are not correlated thus implying that the probability to be attacked is as good as randomly assigned with respect to the orographic conditions. Figure 1 shows that this condition is fulfilled in our data.

[Fig. 1]

[7] An example for the need of these fixed effects is crime rate in the south. After Italian unification southern Italy (i.e. the area that is closer to North Africa) was characterized by widespread peasant revolts (so-called Brigandage) and harsh military repressions until mid-1860s (Accetturo et al., 2016). The need for defense from such attacks might have induced the population to choose more sheltered locations.

Baseline results are provided in Table 1 (see: Table A1 for the descriptive statistics). In Column 1, we provide the result we obtain from estimating equation (1), omitting to control for territorial fixed effects. This means that we the probability of attack actually captures the difference between Southern (high probability) and Northern (low probability) locations. Main effects have the expected sign. The higher the protection granted from a location (rugged location) the lower the population; cities with a high probability of attack are instead larger implying that southern locations were (in 1871) on average larger than Northern ones. Our variable of interest is positive and enters with high significance. In terms of the standardized coefficient is such that a s.d. increase of the interaction term is associated with a 36% rise in the s.d. of the municipal population.[8] Obviously, these correlations might capture aspects related to differences across Italian territories that only by chance happen to be correlated with our variable of interest. First of all, as the port of departure is in northern Africa, all southern regions are more exposed to attacks. The regions of the South, however, differ from their center-northern counterparts for a number of additional reasons. In 1871 – 10 years after the unification of Italy – differences were likely even more pronounced than the current ones (see Felice, 2014), as regions were previously part of different national entities. To control for these aspects Column 2 adds a dummy for southern municipalities. In this case, the coefficient of interest is estimated by using within-area variation only. Therefore, the area-common distance with Tunis is differentiated away. In Column 3 and Column 4 we include increasingly detailed geographic fixed effects: respectively for the 4 Italy's repartitions (NUTS1: North-West, North-East, Centre, and South) and the 19 regions (NUTS2 – there were only 19 regions out of the current 20 because at that time Trentino Alto Adige was not part of the country). In this last specification, which is the most conservative one and it is taken to be our preferred, the region-common distance to Tunis is differentiate away and the probability of being a target of a pirate's attach is measured as deviation from the region average likelihood of being assaulted. According to the specification of Column 4, the standardized coefficient on the interaction term is equal to 0.21 and is highly significant.

[Tab. 1]

[8] The implied magnitude is that a standard deviation increase in our measure of protection determines a rise in population by 39.7% for cities located nearby Naples (in the South, i.e. an exposed area); the same rise in ruggedness implies a rise in population by 26.5% nearby Genua, in a relatively sheltered area.

To the extent that our evidence can correctly be attributed to the attacks of the pirates - occurring from the VIII to the XIX century - the effect of our variable of interest on the spatial distribution of population *before the start of the raids* should be zero (or anyway different from that found for 1871). Column 5 of Table 1 presents a placebo experiment, which is very supportive in this regard. By using historical data from Pleaides database of ancient places, we are able to select among the 7,694 Italian municipalities recorded in 1871, those that were Roman sites. 680 municipalities can be classified as previously Roman spots. As we do not have information on the respective populations, we use as outcome a dummy equal to 1 for the 1871 municipalities with Roman origins and re-estimated eq. 1 by LPM. The results suggest that the need of protection was absent before the VIII century. Our coefficient of interest enters with a negative, rather than positive, sign and it is highly significant. This is not surprising as traffics with northern Africa under the Roman Empire were limited to commercial flows, while the Barbarian invasions of the III century followed a path from north Europe to Italy.

3.1. Robustness with respect to the definition of *Pr(attack)*

As explained in Sect. 2, the majority of the boats raised anchor from Tunis both in the first and the second wave. However, other ports in northern Africa were likely involved in pirate's departures. This is the case of Algiers and Tripoli, according to Kontsam (2016). In Table 2 (Column 1 and Column 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 sailing paths are highly correlated with our baseline path, this check should be taken *cum grano salis*. The results, in any case, nicely survive. We also calculate Pr(attack) 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). Results are depicted in Column 3. During the period in which boats from northern Africa carried pirates, other ships sailing the Mediterranean. They were peaceful vessels, transporting goods and people for trade purposes (they were also targeted by the pirates). The main port of departures for peaceful ships were Barcelona and Marseille. In Column 4 and

Column 5, we measure the probability of receiving an attack by the inverse distance from these two places, respectively. As no pirates casted off from these ports, no effect should be detected. Results from these placebos are very supportive.

[Tab. 2]

3.2. Robustness with respect to the definition of *Protection*

Historical accounts suggest that adequate protective sites were found in inward spaces with a high altitude and on a rugged territory. Our variable Protection therefore considers these three elements jointly through a Principal Component routine. The first component, which we take to proxy for Protection, is the only one with an eigenvalue larger than one and explains 0.57% of the common variance (see Table A2). It is highly correlated with altitude and slope, while the correlation with the dummy for not being on the coast is much lower. As Italy has a very assorted territory a relevant question is whether the three elements have to be jointly considered. For instance, in regions with very high mountains altitude alone could have been enough to discourage attacks. On the same token, very rugged territories could have defended residents even a lower altitude. Table 3 provides the regression results obtained for the baseline specification of Table 1, Column 4 when the single components of our composite index for Protection are used. Except for the dummy for not being on the coast, which however has a low correlation also in the Principal Component exercise, the coefficient on the interaction enters with very high significance and magnitudes that are consistent with those previously found. Note also that our results suggest that ruggedness has a bigger role than altitudes. This makes sense since the attacks were more frequent in the South of Italy and along the Tyrrhenian coast. The inward areas of these shores are featured by mountains of reduced heights.

[Tab. 3]

3.3. Areas differently exposed to the raids

Section 2 describes the exposition of the Italian territories to the raids. The Tyrrhenian coast was massively targeted. The north Adriatic one was instead relatively preserved because of the protection granted by the Republic of Venice, a worldwide naval power. Table 4 presents the results we obtain by splitting the sample between areas targeted and areas mostly unspoiled. As the two areas correspond to the two sides of the Italian peninsula, for this exercise we refer to a bandwidth of 50 kilometres from the coast, to avoid a wrong attribution to untargeted areas of exposed territories. The results for the most affected areas are in Column 1. The regions we include in this experiment are Liguria, Tuscany, Lazio, Campania, Basilicata, Campania and Sicily. The findings for the untargeted areas are presented in Column 2. In this case we include in the estimation sample: Marche, Emilia-Romagna, Veneto and Friuli Venezia Giulia. That is, the coastal areas that relatively sheltered by the Venetian naval power. The results we obtain confirm our insight. No impact is found for untargeted areas, while for the most affected areas the estimated effect is positive and significant.

[Tab. 4]

4. Persistency

Table 5 provides the estimation results we obtain by using as dependent variable municipal population measured at Census dates from 1861 to 2001. The results are striking. The impact of the pirates' attacks is still detectable in 1981 (even though statistically significant up to 1961), more than one century and a half after the raids were terminated. The pattern of point estimates we obtain is monotonically (and slowly) decreasing overtime: for instance, in 1951 the effect is estimated to be a 50% of that measured in 1871. This persistency is noteworthy. 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 XIX century and the twenties of the XX century. Finally, at the beginning of the XX century the country 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 persistency might have been helped by the overall structural characteristics of Italy's economy from the Unification to the aftermath of WWII, mainly agricultural with a reduced need of migrants, and the prolonged period of Fascist dictatorship, which banned migration (see: Andini et al, 2016). The effect we estimate ceases to exist after 1981. This is not surprising as internal migrations, manly south to north and rural to urban, were a distinctive feature of the Italian industrialization process up to middle 1970s.

[Tab. 5]

5. Effects on local economic development

So far we have consistently shown that, due to the fear of attacks by pirates coming from the Northern Africa, population tended to concentrate relatively more in easy-to-protect locations in regions that were more exposed to raids. Unfortunately, highly defensible locations are also generally less suitable for the economic activity. In general, inner locations have a lower market access; high altitude determines a reduction in the availability of crops; sloped terrains are generally characterized by a lower productivity in agriculture and, once again, by a more limited access to external markets.

This might imply that the fear of pirates' attacks have determined an abnormal growth for locations that were not able to sustain such a large population. In other words, people were mislocated with respect to the actual productivity of such areas.

To test this hypothesis we have to check whether economic outcomes in those areas were worse than those registered in other parts of the country. Data availability constrains our choices in terms of indicators and time span. Information on wages and rents, for example, are available only for the last 15 years for Italian cities but overpopulation was registered only until 1981. From 1951 we use information on:

- employment rate
- number of non-agricultural plants per capita
- human capital
- ageing index

- share of employees in agriculture

In order to capture whether overpopulation determined a worse economic performance in the affected areas, we run equation (1) by using, as outcome variables, those listed above. For the sake of simplicity Tables 6 to 10 report the estimated coefficient (and its standardized version) for the interaction term only.

Results for the employment rate (i.e. number of employees over population) are displayed in table 6. Data availability data constrains our analysis for the period 1951-1991 as more recent censuses were not harmonized for the older ones. In 1951 overpopulated localities were characterized by a lower employment rate (-0.10 in terms of standard deviation); the result is still significant in 1961 although it was reduced by half. From 1971 on the negative effect disappears probably due to the large outflow of population from those areas.

[Tab. 6]

Table 7 presents the results for non-agricultural plants per capita using the same dataset of table 6. In 1951 the negative effect of overpopulation on plants per capita was similar to the one for employment rate (-0.10 the standardized coefficient). However, subsequent outmigration determined a more than proportional decrease in the number of plants in the following years (-0.30 the standardized coefficient in 1991); this is compatible with a core periphery pattern with decreasing trade costs as predicted by the New Economic Geography literature (Baldwin et al., 2002).

[Tab. 7]

Outmigration from overpopulated areas also involved a change in the age structure of the staying individuals. Table 8 presents the results by using the Ageing Index (i.e. the share of individuals above 65 years old) as an outcome variable. Still in 1951, age structure in affected municipalities was not, coeteris paribus, different from other cities. Migration flows that started in 1950s however mostly involved younger individuals thus determining a rise in the ageing index by almost 20% in terms of standard deviation.

[Tab. 8]

Sluggish economic conditions in affected areas also reduced the incentives to accumulate human capital. Table 9 presents the results by using the share of individuals with at least a secondary school diploma that is the only outcome variable that is consistently available from 1951 on.[9] Still in 1951 overpopulated localities were characterized by a lower average schooling than other areas (-0.15 in standard deviation terms). The effect is still detectable in 2001, though the magnitude is reduced by one-third probably due to the fact that migration flows in the 1950s and 1960s mostly involved unskilled workers.

[Tab. 9]

Finally, milocation had consequences on the productive structure of local economies. Table 10 presents the estimates when we use the share of agricultural employment as an outcome variable. Estimation results show that overpopulated were relatively more specialized in agriculture and this feature persisted until 2001. The fact that high protection areas are also less suitable for high value-added agricultural production due to altitude and slope of terrain implies that those areas mostly lived on subsistence farming.

[Tab. 10]

6. Suggestive evidence on aggregate effects

Overpopulation due to the fear of pirates' attacks determined that, as we have seen in the previous section, individuals concentrated in areas that were not suitable for economic development. This implied lower employment rates, slower accumulation in human capital, and overspecialization in subsistence agriculture.

[9] Data on tertiary education are available from 1971 on with results much in line with those of table 9 (results available upon request).

Did these features have aggregate effects for the development of the entire country?

The question whether agglomeration fosters aggregate growth has been widely investigated under a theoretical point of view (Baldwin et al., 2001; Fujita and Thisse, 2002; Baldwin et al., 2002; Accetturo, 2010; Fujishima, 2013). Empirical evidence is more scant (Bruelhart and Sbergami, 2009; Gardiner et al., 2010) and supports the idea that the emergence of a core-periphery pattern correlates with a more intense aggregate growth.

In this section we provide some suggestive evidence (i) on the fact that most affected areas were characterized by the lack of primary cities (ii) on the positive relationship between urban concentration and value added per capita.

We use the province (i.e. the equivalent of a US county in terms of size) as the unit of analysis following a long tradition in macroeconomics that uses regions or cities as laboratories to understand sources of differences across countries (for example, Barro and Sala-i-Martin, 1991 and 1992; Gennaioli et al., 2013). We proceed in three steps.

First, we compute an index of dispersion for population within each province. This is calculated by estimating the Zipf coefficient, i.e. the correlation between the (log) rank that each municipality has within the province in terms of population and its (log) population. In formula, for all years, we estimate:

(2) $Log(rank_{ip}) = a_p + b_p Log(POP_{ip}) + e_{ip}$ for all p=1,…,93

b_p is the value of the Zipf coefficient for province p. It should be equal to -1 if the Zipf law is respected; however, since we are working on the entire population of Italian cities, we expect that it is always different from -1 (Eeckhout, 2004). The Zipf coefficient can be interpreted as a measure of concentration of population across different municipalities within each province. The lower the index the more dispersed is population and the less important is the major urban center in the province. Results for the distribution of the Zipf coefficient are presented in the map in Figure 2a for 1951.

[Fig. 2]

In the second step of the analysis we correlate the Zipf coefficient with the importance of the milocation shock in the province. In particular we run the following regression:

$$(3)\ Zipf_p = k_2 + \alpha_2\ Pr(attack)p + \beta_2\ Protection_p + \gamma_2\ (Pr(attack)_p{*}\ Protection_p) + \varepsilon_p$$

where variables indexed by p indicate averages at province level of the variables of equation (1).

We expect γ_2 to be negative and significant, thus implying that the larger the mislocation shock in the province, the more dispersed is population across locations. Graphical evidence of such negative correlation is detectable by comparing panel a and b of figure 2. Results from the estimates of equation (3), for each year, are displayed on table 11 and confirm our intuition that $\gamma_2<0$.

[Tab. 11]

In the third step we check whether the population dispersion induced by pirates' attacks has a negative impact on aggregate average provincial income. To this aim we used the log value added per capita computed by the Tagliacarne Istitute at provincial level from 1951 on as dependent variable for the estimation of the following equation:

$$(4)\ Log(VA\_pc)_p = k_3 + \alpha_3\ Pr(attack)_p + \beta_3\ Protection_p + \gamma_3\ Zipf_p + \varepsilon_p$$

if the city size distribution matters for economic development we should observe $\gamma_3>0$.

For the estimation of equation (4) we use both OLS and 2SLS. In the latter case, we instrument Zipfp with ($Pr(attack)_p{*}\ Protection_p$) that is the population shock that determines a more distribution of population across locations. Results are displayed in table 12 and show that the higher the concentration of population in the province the higher the value added per capita. This is confirmed both in the 2SLS and in the reduced form estimates.[10]

[Tab. 12]

Obviously, the results of this section must be taken *cum grano salis*. The use of more aggregate unit of observation prevents us from using the set of fine geographical

[10] In this case we regress the log value added per capita directly on the instrument.

dummies. This implies that there can be other omitted variable that may be correlated with both the urban development and income levels. Moreover, the magnitude of the effect estimated with 2SLS should be considered with caution. As first stage statistics show, the instrument is rather weak thus implying that estimated effects are biased; however, as Angrist and Pischke (2009) point out, 2SLS results are informative of the sign and the significance of the effects even when the instrument is weak, especially when reduced form estimations are significant and have the right sign.

Overall, these results suggest that mislocation hampered the emergence of primary cities in the province and, as a result, this process did not allow the development of agglomeration economies with positive spillovers on the entire economy.

7. Conclusions

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

These results have relevant consequences from both a positive and normative perspectives.

From the positive side, we have shown the importance of first-nature advantages in shaping the economic outcomes. In the areas affected by pirates' attacks the advantages of overpopulation in terms of agglomeration economies were widely overcome by the disadvantages in terms of productivity; as a results, once the historical event that determined concentration was over, those locations slowly depopulated.

On the normative side, this paper cast some doubts on the economic foundations of many policies for local development aimed at peripheral areas. Many countries have policies aimed at the development of areas characterized by low productivity and weak fundamentals (see Accetturo and de Blasio, 2012, for a review). Italian tradition is particularly strong; quite recently (2012), for example, the Italian government has proposed a project for the development of Internal Areas (Aree Interne) with the aim to resist depopulation and attract economic activities to trigger agglomeration economies.[11] The results of this paper are quite pessimist on the sensibleness of such policies: public resources in those areas are probably not able to overcome the lack of first-nature advantages with possible negative consequences not only on the development not only at local level but also at aggregate one.

[11] See: http://www.dps.tesoro.it/aree_interne/ml.asp

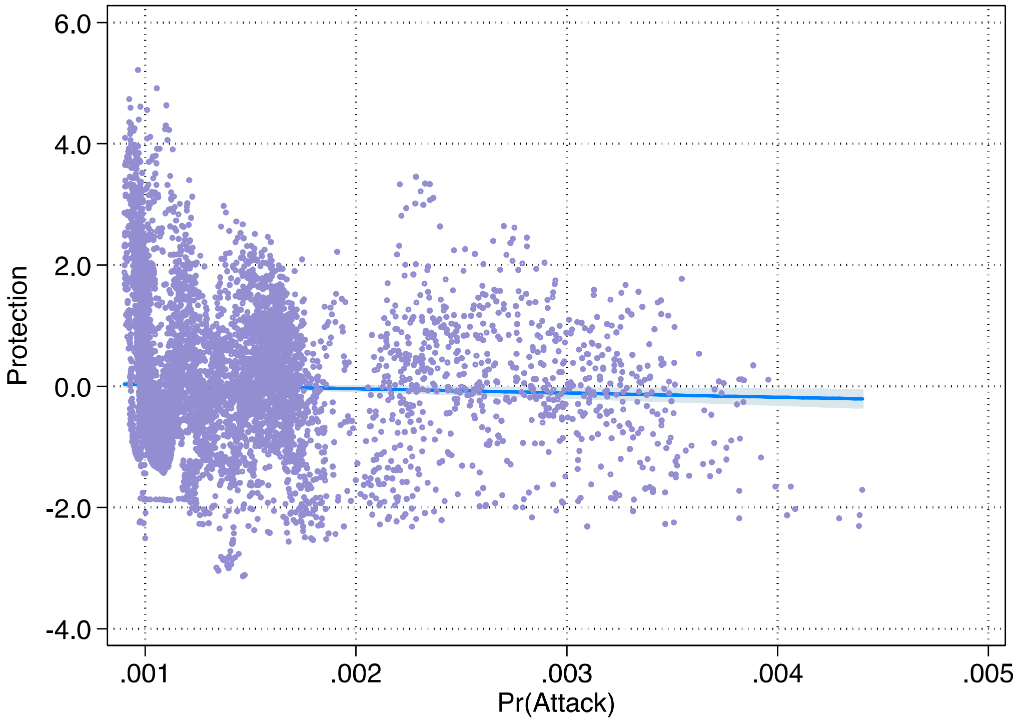
References

- Accetturo, A. (2010). "Agglomeration and growth: the effects of commuting costs." *Papers in Regional Science* 89.1: 173-190.
- Accetturo, A. and G. De Blasio (2012). "Policies for local development: An evaluation of Italy's “Patti Territoriali”." *Regional Science and Urban Economics* 42.1: 15-26.
- Accetturo, A., A. Linarello, and A. Petrella (2015). "Legal Enforcement and Global Value Chains: Micro-Evidence from Italian Manufacturing Firms." Bank of Italy, mimeo.
- Accetturo, A, M. Bugamelli, and A. Lamorgese (2016). " Law Enforcement and Political Participation: Italy, 1861-65", Bank of Italy, mimeo.
- 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.
- 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., J.-S. Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton university press.
- 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.
- Bleakley, H. and J. Lin. (2012) "Portage and path dependence." *The quarterly journal of economics* 127.2: 587.
- 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.
- Colley, L. (2004) *Captives: Britain, Empire, and the World, 1600–1850*, Anchor Books Edition, New York.
- 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.
- 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*.
- Gennaioli, N., R. LaPorta, F. Lopez-de-Silanes, and A. Shleifer (2014). "Growth in regions." *Journal of Economic growth* 19.3: 259-309.
- Gosse, Philip (1933). *The history of piracy*. Tudor Publishing.
- 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.
- Kline, P. and E. Moretti (2014). "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics* 129.1: 275-331.
- Konstam, A. (2016) *The Barbary Pirates 15th-17th Centuries*. Bloomsbury Publishing.
- Michaels, G. and F. Rauch (2016). *Resetting the Urban Network: 117-2012*. Forthcoming Economic Journal.
- Nunn, Nathan, and Diego Puga. "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics* 94.1 (2012): 20-36.
- 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.
- 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.

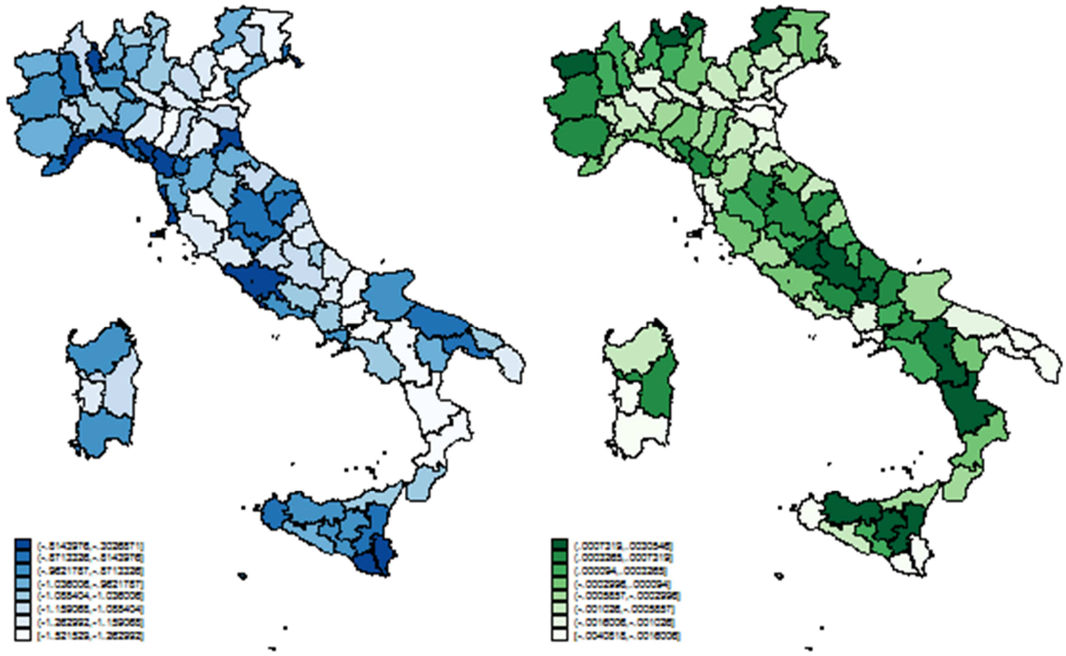
Figure 1. Correlation between Pr(Attack) and Protection



Notes: [xxxxxxxxxxxxxx]

Figure 2. Population shock and city size distribution

(a) Zipf coefficient (b) Population shock



Notes: [xxxxxxxxxxxxxx]

Table 1. The impact of pirate attacks on population, baseline estimates

Dependent variable:	Log population in 1871				Roman settlement
	(1)	(2)	(3)	(4)	(5)
Protection	-0.3427***	-0.3311***	-0.2609***	-0.2230***	0.0178**
	[0.0252]	[0.0256]	[0.0258]	[0.0228]	[0.0081]
Stand. coef.	*-0.4764****	*-0.4604****	*-0.3628****	*-0.3100****	*0.0783***
Pr(attack)	194.3953***	73.7128**	-57.9968*	391.8278***	2.6971
	[21.3565]	[30.2601]	[33.0529]	[65.9459]	[25.7114]
Stand. coef.	*0.1167****	*0.0442***	*-0.0348**	*0.2352****	*0.0051*
Protecion*Pr(attack)	183.4539***	171.4947***	130.2605***	106.7621***	-30.6865***
	[19.5162]	[19.8421]	[20.2571]	[17.3262]	[6.8180]
Stand. coef.	*0.3605****	*0.3370****	*0.2560****	*0.2098****	*-0.1909****
Constant	7.3815***	7.6747***	7.3775***	6.9471***	0.0195
	[0.0301]	[0.0571]	[0.0375]	[0.0739]	[0.0277]
Area Dummies	NO	2	4	19	19
R^2	0.045	0.050	0.130	0.226	0.086
No. Obs.	7694	7694	7694	7694	7694

Notes: OLS regressions. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. The standardized coefficient is calculated as xxxx. Protection is the first PC, as described in Table A2. Pr(attack) is taken to be the geodetic distance from Tunis.

Table 2. The impact of pirate attacks on population, robustness w.r.t. Pr(attack)

Dependent variable: Log population in 1871	(1) Algiers	(2) Tripoli	(3) Weighted Avg.	(4) Barcelona	(5) Marseille
Protection	-0.3000***	-0.2371***	-0.2393***	-0.0233	-0.0571***
	[0.0553]	[0.0261]	[0.0248]	[0.0378]	[0.0207]
Stand. coef.	*-0.4171****	*-0.3296****	*-0.3328****	*-0.0323*	*-0.0794****
Pr(attack)	1081.1267***	972.3510***	493.3865***	291.6485**	82.1818***
	[226.1758]	[225.2001]	[78.2826]	[125.0100]	[31.7117]
Stand. coef.	*0.2034****	*0.3247****	*0.2483****	*0.0932***	*0.0837****
Protecion*Pr(attack)	235.6542***	176.9835***	129.4592***	-50.8234*	-13.9975
	[60.1555]	[29.4974]	[20.4979]	[29.6339]	[8.5158]
Stand. coef.	*0.3082****	*0.2278****	*0.2324****	*-0.0907**	*-0.0473*
Constant	6.2901***	6.6843***	6.8639***	6.9009***	7.0805***
	[0.2233]	[0.1596]	[0.0826]	[0.1983]	[0.1098]
Area Dummies	19	19	19	19	19
R^2	0.222	0.226	0.227	0.219	0.219
No. Obs.	7695	7695	7695	7695	7695

Notes: OLS regressions. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. The standardized coefficient is calculated as xxxx. Protection is the first PC, as described in Table A2. Pr(attack) is taken to be the geodetic distance from the localities described in the column.

Table 3. The impact of pirate attacks on population, robustness w.r.t. Protection

Dependent variable: Log population in 1871	Protection: individual components		
	(1) Non coastal	(2) Altitude	(3) Slope
Protection	-0.4141***	-0.0014***	-0.0005***
	[0.1372]	[0.0001]	[0.0000]
Stand. coef.	*-0.1284****	*-0.4390****	*-0.3251****
Prob(attack)	301.2007***	163.4232**	153.7602**
	[78.4094]	[67.6049]	[66.3581]
Stand. coef.	*0.1808****	*0.0981***	*0.0923***
Protection*Prob(attack)	74.9473	0.5961***	0.3876***
	[68.0561]	[0.0730]	[0.0387]
Stand. coef.	*0.0506*	*0.2982****	0.3798***
Constant	7.3556***	7.4805***	7.2155***
	[0.1371]	[0.0805]	[0.0768]
Area Dummies	19	19	19
R^2	0.217	0.245	0.222
No. Obs.	7694	7694	7694

Notes: OLS regressions. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. The standardized coefficient is calculated as xxxx. Pr(attack) is taken to be the geodetic distance from Tunis. Protection is taken to be the single component of the composite index, as reported in the column.

Table 4. The impact of pirate attacks on population, most affected vs. least affected areas

Dependent variable: Log population in 1871	Most affected	Least affected
Protection	-0.1397**	-0.3831
	[0.0607]	[0.4303]
Stand. coef.	*-0.1599***	*-0.4646*
Prob(attack)	307.6725***	-594.4951
	[69.5958]	[840.8092]
Stand. coef.	*0.2106****	*-0.1070*
Protection*Prob(attack)	66.2478**	231.9297
	[29.2565]	[353.0705]
Stand. coef.	*0.1603***	*0.3217*
Constant	7.1233***	8.2595***
	[0.1290]	[0.8193]
Area Dummies	7	4
R^2	0.194	0.073
No. Obs.	2278	623

Notes: OLS regressions. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. The standardized coefficient is calculated as xxxx. Protection is the first PC, as described in Table A2. Pr(attack) is taken to be the geodetic distance from Tunis. Most affected regions (Tyrrhenian coast): Liguria, Tuscany, Lazio, Campania, Basilicata, Campania and Sicily. Least affected regions (Adriatic coast): Marche, Emilia-Romagna, Veneto and Friuli Venezia Giulia.

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

Dependent variable: Log population in…	Protection*Pr(attack)	Standard error	Stand. coef.
1861	96.4568***	[18.0268]	0.1924***
1871	106.7621***	[17.3262]	0.2098***
1881	98.0236***	[17.1862]	0.1914***
1901	86.9329***	[17.4878]	0.1657***
1911	84.0760***	[17.7355]	0.1575***
1921	79.4488***	[18.0718]	0.1450***
1931	75.5992***	[17.3118]	0.1341***
1936	76.4090***	[17.2795]	0.1331***
1951	58.3745***	[17.6795]	0.0976***
1961	37.8397**	[18.5151]	0.0602**
1971	22.7385	[19.8769]	0.0335
1981	10.6886	[20.2412]	0.0150
1991	-0.4298	[20.9411]	-0.0006
2001	-3.3559	[21.3846]	-0.0045

Notes: OLS regressions. Robust standard errors are in parenthesis. * (**) [***] denotes significance at the 10% (5%) [1%] level. The standardized coefficient is calculated as xxxx. Protection is the first PC, as described in Table A2. Pr(attack) is taken to be the geodetic distance from Tunis.

Table 6. Employment rate

Dependent variable: Employment rate	Interaction	Standard error	Standardized coef.
1951	-6.2302***	[1.4392]	-0.1033***
1961	-21.5544***	[4.1080]	-0.0628***
1971	0.0385	[2.2779]	0.0004
1981	-4.0837*	[2.4553]	-0.0429*
1991	-0.8234	[2.4171]	-0.0089

Notes:

Table 7. Plants per capita

Dependent variable: Plants per capita	Interaction	Standard error	Standardized coef.
1951	-0.7410***	[0.2310]	-0.1031***
1961	-3.8477***	[0.7117]	-0.0936***
1971	-1.8293***	[0.3312]	-0.1824***
1981	-3.4947***	[0.4466]	-0.2303***
1991	-4.5797***	[0.4966]	-0.3042***

Notes:

Table 8. Ageing Index

Dependent variable: Ageing rate	Interaction	Standard error	Standardized coef.
1951	-95.5111	[58.2336]	-0.0397
1961	281.5362***	[73.9847]	0.0868***
1971	1013.4915***	[118.2591]	0.2070***
1981	1202.7498***	[158.3471]	0.1813***
1991	1234.6284***	[174.8583]	0.1720***
2001	1610.8365***	[197.3213]	0.2027***

Notes:

Table 9. Human Capital

Dependent variable: Share of individuals with at least a secondary school diploma degree	Interaction	Standard error	Standardized coef.
1951	-109.0750***	[25.9026]	-0.1481***
1961	-119.6492***	[31.2867]	-0.1357***
1971	-203.2106***	[46.5931]	-0.1527***
1981	-205.1712***	[69.9146]	-0.0967***
1991	-99.0289	[100.0929]	-0.0314
2001	-380.8209***	[119.0490]	-0.1038***

Notes:

Table 10. Share of Agricultural employment

Dependent variable: Share of agricultural empl.	Interaction	Standard error	Standardized coef.
1951	5066.8937***	[387.5294]	0.3605***
1961	4073.4357***	[352.8236]	0.3062***
1971	3534.5842***	[302.7134]	0.3044***
1981	4505.5130***	[395.4015]	0.4148***
1991	2640.1201***	[236.2047]	0.3716***
2001	1776.5709***	[180.5214]	0.3367***

Notes:

Table 11. Effects of Mislocation on the spatial concentration of population

Dependent variable: Zipf coefficient in…	1951	1961	1971	1981	1991
Protection	0.0606	0.0611	0.0792*	0.0936**	0.1131***
	[0.0518]	[0.0460]	[0.0405]	[0.0387]	[0.0401]
Stand. coef.	*0.2514*	*0.2735*	*0.3910**	*0.4941***	*0.6181****
Pr(attack)	13.9127	-4.4353	-25.9855	-33.9208	-36.7887
	[30.3243]	[27.8357]	[25.2605]	[23.4011]	[22.5575]
Stand. coef.	*0.0414*	*-0.0143*	*-0.0921*	*-0.1286*	*-0.1444*
Protecion*Pr(attack)	-68.0965*	-72.3874**	-83.2063***	-83.1973***	-89.1879***
	[36.9482]	[32.7512]	[27.7632]	[25.4716]	[25.2841]
Stand. coef.	*-0.3925**	*-0.4503***	*-0.5705****	*-0.6104****	*-0.6774****
Constant	-1.0605***	-0.9703***	-0.8559***	-0.8017***	-0.7732***
	[0.0492]	[0.0450]	[0.0408]	[0.0384]	[0.0376]
R^2	0.044	0.050	0.059	0.052	0.058
No. Obs.	93	93	93	93	93

Notes:

Table 12. Effects of Mislocation on aggregate income

Dependent variable: Log value added per capita in…	1951	1961	1971	1981	1991
			(a) OLS		
Zipf coefficient	0.2273*	0.4122***	0.2305*	0.3699***	0.2307*
	[0.1261]	[0.1435]	[0.1241]	[0.1213]	[0.1278]
Protection	0.0092	-0.0120	-0.0227	-0.0213	-0.0242
	[0.0407]	[0.0347]	[0.0274]	[0.0236]	[0.0243]
Pr(attack)	-296.8194***	-353.9425***	-227.9798***	-271.1456***	-275.9186***
	[45.5039]	[42.1774]	[31.9441]	[29.9261]	[29.9542]
Constant	1.9659***	2.6900***	2.7705***	3.2326***	3.3022***
	[0.1492]	[0.1489]	[0.1207]	[0.1168]	[0.1118]
R^2	0.287	0.397	0.311	0.452	0.435
No. Obs.	93	93	93	93	93
			(b) IV		
Zipf coefficient	2.3778*	2.3595***	1.2546**	1.6638***	1.2446***
	[1.3473]	[0.9024]	[0.5357]	[0.4740]	[0.4793]
Protection	0.0650	0.0483	0.0045	-0.0055	-0.0239
	[0.0666]	[0.0551]	[0.0342]	[0.0332]	[0.0299]
Pr(attack)	-379.6237***	-396.2108***	-232.1406***	-266.1312***	-271.2750***
	[83.8690]	[59.5810]	[39.3137]	[38.8237]	[33.1598]
Constant	4.3183***	4.6485***	3.6887***	4.3227***	4.1304***
	[1.4685]	[0.8887]	[0.4698]	[0.3886]	[0.3911]
First stage	-68.09647*	-72.3874**	-83.20629**	-83.19733**	-89.1879**
	[36.94821]	[32.75123]	[27.76318]	[25.47163]	[25.2841]
F-first stage	3.40	4.89	8.98	10.67	12.44
Reduced Form	-161.9183***	-170.8008***	-104.3884**	-138.4199***	-111.0026***
	[53.0205]	[48.1235]	[43.4093]	[26.9513]	[31.9975]
R^2	-1.360	-0.646	-0.135	-0.165	0.078
No. Obs.	93	93	93	93	93

Notes:

TABLE A1

DESCRIPTIVE STATISTICS

	Mean	Standard Deviation
Ruggedness:		
Non-coastal (dummy 0-1)	0.917	0.274
Altitude (continuous, in meters)	341.778	283.040
Slope (continuous, in meters)	616.385	618.614
Inverse distance from Tunis (continuous, 1/km)	0.001	0.001
Roman settlement (dummy 0-1)	0.085	0.280
Log(population-1871) (continuous)	7.655	0.886
Log(population-1951) (continuous)	7.931	1.042

TABLE A2

PRINCIPAL COMPONENTS RESULTS

	PC1	PC2	PC3
Scoring coefficients:			
Non-coastal	0.281	0.937	0.206
Altitude	0.698	-0.052	-0.713
Slope	0.658	-0.345	0.669
Eigenvalues	1.695	0.969	0.335
Explained variance	0.565	0.323	0.111