



The economic complexity of the Roman Empire as inferred from Latin inscriptions



Matteo Mazzamurro^{1,3}, Petra Heřmánková^{1,3}, Michele Coscia²✉ & Tom Brughmans²

Economic complexity is a powerful tool to estimate the productive capabilities and future growth of modern economies. Little is known about how economic complexity evolves over long periods in history. In this paper, we use archeological evidence from the Roman Empire in the form of short texts preserved on a durable material (i.e., inscriptions) to estimate the economic complexity of the various provinces of the empire. By connecting the occupations listed in the text of inscriptions with the location in which the inscribed objects were found, we can estimate that the most complex areas during the first four centuries of the Roman Empire have a remarkable and statistically significant overlap with the most complex countries today. This is a surprising result running counter the expectations from historiography, and we lack an explanation for the reason of this supposed preservation of economic complexity through the ages – both factors pointing at the need of future work to integrate Greek inscriptions. However, this evidence provides a suggestion about how difficult the development of economic capabilities might be.

Economic complexity has emerged in the literature as a way to estimate a country's productive knowledge^{1,2}. In general terms, a country's economy is defined to be “complex” if the country is able to perform a diverse set of economic activities requiring unrelated knowledge more than it would be expected given its population. Normally estimated as the Economic Complexity Index (ECI)—of which there are multiple alternative definitions^{3–5}—it has gained relevance because it is able to predict not only economic growth⁶, but also development trajectories—which economic activities will appear where next^{7–9}. Other applications involve discussing income inequality¹⁰, the green transition¹¹, and more¹².

Given how much of our society's current and future challenges can be informed by economic complexity, we should devote some attention to understanding how we can affect it—specifically how it evolves. While most of the cited works—and more¹³—study ECI changes in the short and medium term, no one has had the opportunity to do so on long historical time scales.

In this paper, we propose to take a long-term – millennia-wide – look at economic complexity. Our effort starts with the use of archeological data on the Roman Empire to create the most accurate picture to date of its occupational and economic structure. This is an interdisciplinary effort, putting together network science and archeology—a novel approach that has been gaining increasing attention in recent years^{14–17}.

The performance of the economy of the Mediterranean world under Roman imperial rule has been studied through the quantitative analysis integrating archeological evidence, such as pottery, coins, and most recently,

inscriptions, with economic analysis^{18–21}. Recent studies on the persistence of Roman public road infrastructure into modern times^{22–24} provide crucial evidence that infrastructure networks shaped settlement patterns and economic geography over millennia, suggesting that spatial distributions of occupations documented epigraphically in our dataset reflect deep structural economic patterns rather than ephemeral arrangements.

We build on previous work that reconstructed the occupational profile of the Roman Empire from inscriptions²⁵. We use the largest digital dataset of Latin inscriptions with over 500,000 records as our main source of evidence^{26,27}. These inscriptions contain mentions of ancient occupations, as well as spatiotemporal data to build the occupational profile of the Roman Empire. The reconstruction of the economic structure of the Roman Empire allows us to estimate the economic complexity of the empire's provinces, identifying which areas of the empire contributed most to its complexity.

Finally, we can connect the ancient economic complexity to the modern one. The Roman Empire was a continental polity, that spanned 36 modern countries. We have an ECI estimation for each of these modern countries based on contemporary trade data and we can correlate it with the ECI based on archeological data. We find that there is a weak but marginally significant correlation between ancient and modern ECI in the countries included in the Roman Empire. This is a remarkable connection between two economic structures separated by almost two millennia of history.

We are unable to discern whether this correlation is due to long-term effects of economic complexity or to exogenous factors such as geographical positioning or natural resources²⁸. Moreover, this result runs counter

¹Social Resilience Lab, Aarhus University, Aarhus, Denmark. ²IT University of Copenhagen, Copenhagen, Denmark. ³These authors contributed equally: Matteo Mazzamurro, Petra Heřmánková. ✉ e-mail: mcos@itu.dk

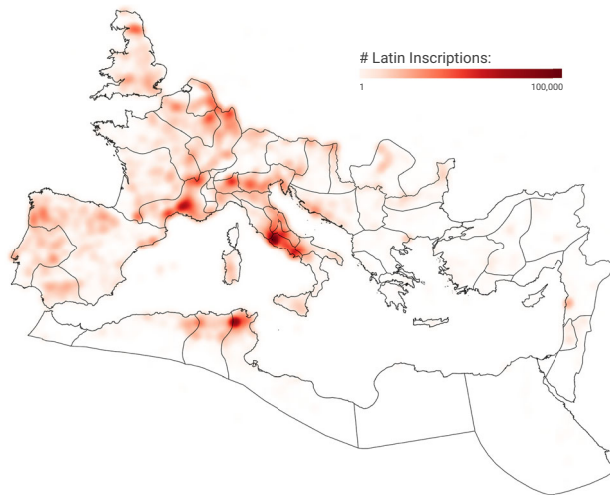


Fig. 1 | The total number of Latin inscriptions in our dataset. Heatmap, the darker the red hue, the more inscriptions were found in the area. Figure generated with QGIS.

expectations from historiography: the assumption is that complexity is related to wealth and urbanization, factors that were more uniformly present in the Eastern Roman Empire than in the Western. However, today the areas that were under the Eastern Roman Empire have lower ECI than the areas that were under the Western Roman Empire. Our work has excluded the vast majority of eastern provinces because we argue the available data did not enable their inclusion. This points at the need to integrate data coming from non-Latin inscriptions to expand our analysis to parts of the empire with different epigraphic traditions. We further emphasize that this result should be understood as almost exclusively reflecting the situation in the former Western Roman Empire. Nevertheless, this result is the best picture of Roman economic complexity we can produce as of today, and it hints at potential implications for development policies.

Regardless of the causes, relative economic complexity rankings appear to be enduring, meaning that—without a conscious and possibly global effort—the income disparities between developed and developing countries might linger indefinitely, as these might be fostered by enduring differences in economic complexity²⁹.

Results

Occupational structure

We analyze the digitized textual content of 525, 870 Latin inscriptions created during the Roman Empire. Figure 1 shows the geographical coverage of these inscriptions.

Most inscriptions date to the first four centuries CE²⁶, a period that this study focuses on as it covers most of the duration of the Latin-speaking western Roman Empire. Inscriptions are short texts surviving on an object made from a durable material and were created for a variety of purposes: to mark someone's grave and provide details of their life, to celebrate their life achievements and bestow honors, to regulate, to inform, but also to materialize an insult, or joke, or to invoke magical powers. Upon computer-assisted reading of these texts, we discovered 8 475 inscriptions that mention at least one occupation (from one to 30 mentions of occupations in a single inscription). Of these, 37% (3 116) are funerary inscriptions, reporting on the deceased person's profession³⁰, and 11.4% are honorific texts (952) typically listing the life achievements of the honouree, including their occupation, known as the *cursus honorum*³¹.

Most inscriptions have a known findspot identified with latitude and longitude coordinates, which allows us to precisely geolocate 10, 353 occurrences of 440 unique occupations throughout the Empire – under the assumption that mobility during the Roman Empire was less frequent than nowadays³². Detailed evaluation of ancient mobility using epigraphic

sources has shown that most people migrated locally, with only a small group of highly mobile individuals³³. We provide more information about the raw data in the Materials & Methods section.

We count the number of occurrences of a given occupation per Roman province. We use the Roman province borders from ca. 200 CE digitized by *The Ancient World Mapping Center*³⁴ and edited by Adam Pažout in 2023. We exclude provinces east of the Euphrates, which were only occupied for a very short period, and therefore, we assume, they did not contribute to the overall economic complexity of the Roman Empire. All boundaries were clipped to modern coastlines. Inclusion of Italian *regiones* enhances the granularity of the results in the Italian peninsula. Unfortunately, a similar finer division is not entirely possible for other provinces. We only include provinces with at least 1 000 inscriptions in Latin to ensure the data is representative of the past provincial situation.

Table 1 reports selected summary statistics about our data: focusing on the first two columns, we show how many inscriptions we find in each province containing at least the mention of one occupation and the number of distinct occupations mentioned in the inscriptions.

For our main results, we use only Latin inscriptions. Latin was one of the principal epigraphic languages of the Roman Empire and survives in large quantities, providing the most datapoints in digitized and accessible format. As shown in Fig. 1, the majority of Latin inscriptions originate from the Western parts of the Roman Empire. Greek inscriptions³⁵, which come predominantly from the Eastern part of the Empire, are treated as supplementary data (see Methods section). We exclude Greek inscriptions from our main analysis due to structural differences in the epigraphic corpus and the unsuitability of current distant reading methods for the peculiarities of Greek epigraphic language, which prevent the application of the same analytical approaches used for Latin texts. Nevertheless, Supplementary Materials Section 2 demonstrates that including Greek inscriptions with certain assumptions about their content yields results similar to those presented here.

Given that we work with archeological (i.e., epigraphic) data, there are a number of biases that would lead these counts not to be accurate. We identify several sources of biases—language of publication, gender, status, and amount of research dedicated to different areas. We detail how we dealt with these sources of bias in the Supplementary Materials Section 1.

Our bias correction results in a probabilistic count of the number of occupations in a given Roman province. The third column of Table 1 shows the total number of people with an occupation we estimate per Roman province.

Moving from province-aggregated counts to counts by province-occupation pair, Fig. 2 shows the most accurate to date representation of the occupational structure of the Roman Empire, counting how many people were involved in a given occupation in a given province. It takes the form of a M_{po} matrix, with p being a province and o being an occupation. Given the high skewedness in the occupation counts, we always operate on the logarithm of the values in M_{po} – for both visualization and analysis, as is standard practice in economic complexity literature¹.

M_{po} has a clear nested structure. Nestedness is a phenomenon originally observed in ecosystems^{36, 37}, by observing which species appear in which ecosystem. In economics, one can count which economic activity appears in which geographical region and the same pattern emerges^{1, 2, 38–40}.

Nestedness is a product of a hierarchy in occupation complexity. In practice, every occupation relies on the simplest one – farmers – to produce the food allowing the other workers to survive. At the same time, the provinces with the lowest complexity employ nothing more than farmers. Therefore, farmers are present in every province. The second least complex occupation is present in all but the simplest provinces. The pattern continues until we get to Latium: a province that can employ the most complex occupations, because it employs all other occupations necessary to support them. Such occupations are found nowhere else in the empire, creating the nested pattern.

Nestedness can be estimated by comparing the M_{po} density on the left and on the right from its isocline—the line dividing the matrix into its dense

Table 1 | Summary statistics about the provinces in our data: number of inscriptions found mentioning at least one occupation, count of distinct occupations in the inscriptions, total count of individuals performing an occupation after our bias corrections, and Economic Complexity Index value, after bias corrections

Province	# Inscriptions	# Distinct Occupations	Occupation Count	ECI
Latium et Campania (Regio I)	4748	337	1,980,688	6.11
Africa Proconsularis	430	70	2,456,283	− 0.02
Venetia et Histria (Regio X)	429	91	1,948,588	0.15
Narbonensis	373	101	1,375,953	− 0.25
Numidia	291	56	846,955	− 0.12
Umbria (Regio VI)	232	55	495,043	− 0.15
Germania Superior	229	66	1,299,113	− 0.26
Etruria (Regio VII)	221	48	1,134,353	− 0.10
Apulia et Calabria (Regio II)	191	65	1,007,767	0.01
Baetica	185	56	1,178,417	0.18
Dacia	182	28	1,231,403	− 0.26
Lugdunensis	181	55	2,312,849	− 0.43
Transpadana (Regio XI)	168	47	697,671	− 0.19
Aemilia (Regio VIII)	159	44	847,233	− 0.16
Samnium (Regio IV)	157	52	490,521	0.05
Pannonia Inferior	154	31	343,667	− 0.21
Britannia	142	39	1,750,000	− 0.30
Hispania Citerior	141	47	5,183,518	− 0.08
Dalmatia	121	36	1,073,787	− 0.23
Pannonia Superior	107	35	806,416	− 0.33
Mauretania Caesariensis	90	33	1,375,740	− 0.01
Picenum (Regio V)	89	31	263,329	− 0.16
Moesia Inferior	88	27	715,209	− 0.21
Belgica	80	38	1,255,436	− 0.40
Asia	78	30	2,750,713	− 0.18
Germania Inferior	77	35	584,992	− 0.37
Liguria (Regio IX)	72	22	483,082	− 0.17
Aquitania	70	34	2,334,436	− 0.34
Lusitania	65	30	1,638,066	− 0.05
Noricum	60	23	684,037	− 0.20
Raetia	60	24	1,039,475	− 0.39
Achaia	48	10	954,313	− 0.01
Macedonia	47	20	744,327	− 0.10
Galatia	46	13	1,727,885	− 0.10
Bruttium et Lucania (Regio III)	39	17	926,888	− 0.13
Moesia Superior	32	22	645,482	− 0.18
Mauretania Tingitana	27	9	334,286	− 0.11
Sardinia	27	18	795,312	− 0.09
Sicilia	26	13	1,140,916	− 0.09
Syria	24	9	2,920,114	− 0.10

upper-triangular part and the rest of the matrix. The bias-corrected M_{po} is more than five times denser on the left of the isocline than on the right. More precisely, it is 5.7 times denser, a value that is statistically significant ($p < 0.01$) when compared to 1 000 randomizations of M_{po} where we generate digital twins of each Roman province (preserving its number of occupations and occupation diversity).

When using HISCO occupations^{41,42} rather than our own occupation classification, the pattern stays: the matrix is 3 times denser to the left of the isocline ($p < 0.01$). Supplementary Material Section 3 contains details on the calculation of the isocline and the comparison with a null model.

Economic complexity

We can exploit the nestedness of M_{po} to calculate the Economic Complexity Index (ECI) of the provinces of the Roman Empire. We use the classical definition of ECI, where a province is ranked as highly complex if it possesses occupations that are only rarely present in the empire. If a province only hosts occupations that are present in many other provinces, it instead has low complexity. We detail the computation of ECI in the Materials & Methods section, as well as an argument as to why we do not consider alternative definitions of ECI⁴.



Fig. 2 | The occupational structure of the Roman Empire. The rows of the matrix represent provinces. The columns of the matrix represent occupations. Each cell reports the count of estimated individuals performing an occupation (column) in a province (row) -- after bias correction. The darker the red hue, the more occupations

were in a given province, using a logarithmic scale. The rows and the columns are sorted by their respective sums.

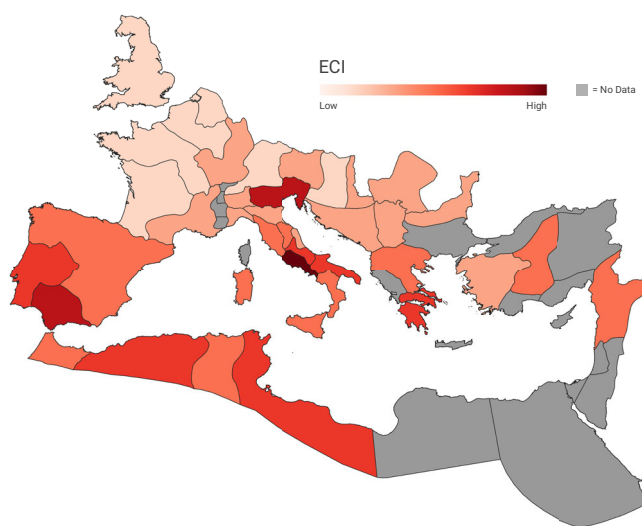


Fig. 3 | The ECI values per province. Provinces in darker color have higher ECIs. Gray color represents provinces without an ECI estimate, due to the low density of surviving datapoints. Figure generated with QGIS.

Table 2 | The top ten most complex modern countries out of the 23 countries with epigraphic coverage, according to the ECI calculation based on epigraphic data (left) and on modern trade (right)

ECI Rank	Ancient	Modern
1	Italy	Germany
2	France	Switzerland
3	Germany	Austria
4	Spain	United Kingdom
5	Belgium	Slovenia
6	Algeria	Hungary
7	Slovenia	Italy
8	Hungary	France
9	Serbia	Belgium
10	Switzerland	Spain

In bold, we align the disagreements in the top ten.

Figure 3 shows the result of ECI per province after bias corrections. As expected, Latium and Rome have by far the highest economic complexity, as the central and most important area of the Roman Empire. We do not find a correlation between ECI and urbanization level across the empire in general; instead, the economic complexity is higher in provinces containing key long-distance trading nodes such as Venetia, Carthage, and Baetica. This suggests that trading and exchanges were important to create a strong ecosystem of occupations.

The ECI rankings we obtain are robust when we use HISCO occupations instead of our own (Spearman rank correlation with $p < 0.001$).

ECI temporal stability

We now turn to study the stability of the Economic Complexity rankings over the course of Western history. Note that we take a correlational approach: we do not investigate what causes an ECI ranking, we only look at its relative stability over the course of a millennium and a half. Establishing whether this stability is endogenous – complexity preserves itself – or exogenous – complexity comes from external factors such as the geographical position of a trade node—is left as a future work.

We calculate the Spearman rank correlation coefficient between the ECI rankings of modern countries both using our ancient epigraphic data and data coming from UN Comtrade for the period 1962–2022⁴³. We use a rank correlation as the specific ECI values are unimportant and the distribution of the values from the Latin inscriptions is skewed, with Italy being a far outlier that could drive the correlation.

We obtain a correlation of 0.439 with $p < 0.05$, indicating that there is a significant positive correlation between the complexities of a country when calculated with data from the Roman Empire and data from modern international trade. Such correlation is mildly stronger (0.522 with $p = 0.011$) when using HISCO occupations.

The fact that we are able to find a significant correlation with only 23 observations is remarkable, considering that the two datasets are 1500 years apart. This result is robust to the introduction of the Greek inscriptions, as we show in Supplementary Material Section 2.

Table 2 shows the top ten countries with the highest ECI values in both the ancient and the modern data. As we can see, there is substantial agreement: 8 out of 10 countries are present in both top tens. Algeria and Serbia were more prominent in ancient times, while data from the Roman Empire failed to capture the rise of the United Kingdom and Austria over the centuries.

Province clusters

Following the literature on economic complexity^{6,44}, in Fig. 4 we build the Occupation Space of the Roman Empire, connecting occupations if they co-appear in the same provinces a significant amount of times—more details in

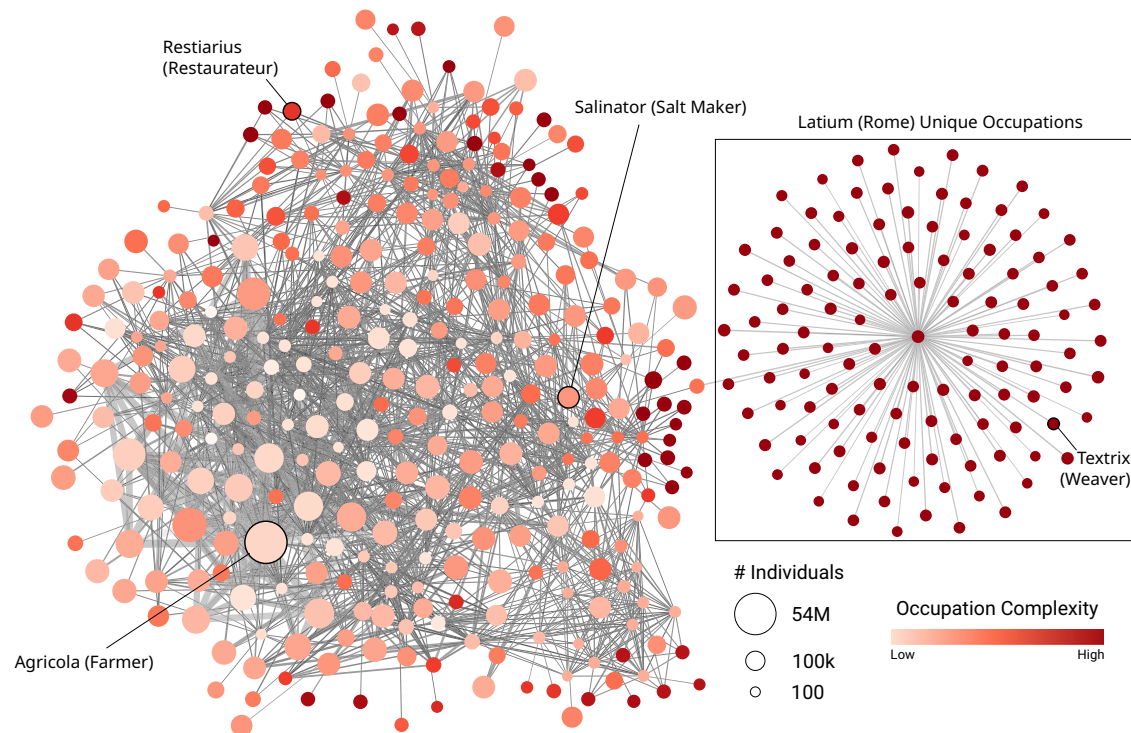


Fig. 4 | The occupation space of the Roman Empire. Each node represents an occupation. Two occupations are connected if they co-appear in significant numbers in the same provinces. The node size is proportional to the number of people in that occupation. The node color is proportional to the complexity of the occupation, from dark red (complex occupation) to white (simple occupation). The edge thickness is proportional to the number of co-appearances, while color and transparency are proportional to the statistical significance. As examples, we chose to

display *agricola* (farmer) as a simpler yet well-represented occupation, *salinator* (salt maker), and *restiarius* (rope maker/seller), more complex occupation with fewer representatives. The Latium cluster (in a box) contains a set of unique, highly complex occupations that do not occur anywhere else. Figure generated with Cytoscape.

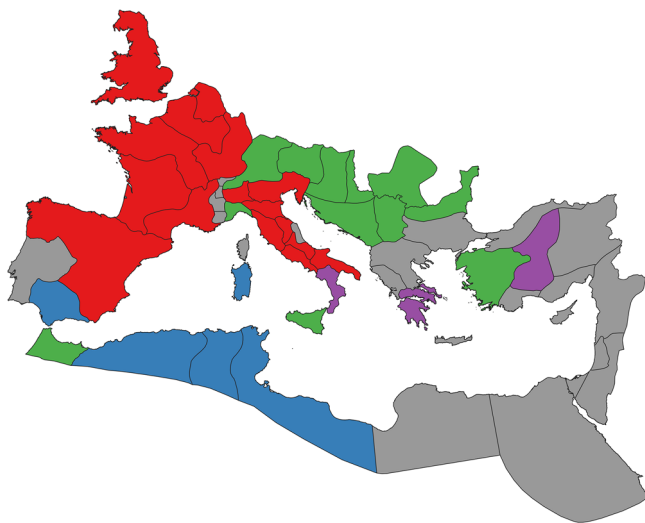


Fig. 5 | The province clusters on the Occupation Space. Provinces in the same color belong to the same cluster. In gray, we have provinces with either no occupation data or outliers that cannot be clustered in clusters larger than three elements. Figure generated with QGIS.

the Materials & Methods section. Such co-appearance is a strong indication of both horizontal and vertical economic integration between the occupations—they either require similar skills or are part of interconnected supply chains.

The Occupation Space has some evident features. The large star on the right includes occupations only found in Latium (the densely populated area around the capital, Rome). The denser main part of the Occupation Space has a noticeable gradient, moving from least to most complex occupation from the core outwards, much like the modern Product Space⁶. The internal structure, just like the Product Space, shows some clusters, with a main core on the left including common occupations such as *agricola* (farmer), *curator* (manager), and *faber tignarius* (carpenter); and smaller communities on the top and bottom right.

We can exploit the structure of the Occupation Space to estimate the similarity between the Roman provinces. Rather than simply counting the number of occupations they have in common, we can use network distances to weight common occupations according to how close they are to each other in the Occupation Space⁴⁵—more details in Supplementary Materials Section 4.

Once we have a province-province distance measure on the Occupation Space, we can analyze all pairwise distances to find clusters of provinces⁴⁶—which we show in Fig. 5. Provinces are grouped together if they have a similar prevalence of occupations that are nearby in the Occupation Space. The red cluster represents the core of the Western Empire; the blue cluster puts together mostly North African provinces; the green cluster demonstrates that the provinces along the Danube had more similarity in the documented occupations due to the enhanced military presence in the region. We describe the clustering procedure in more detail in the Supplementary Materials Section 4.

Discussion

In this paper, we use digitized geolocated inscription data to create the most accurate picture to date of the occupational structure of the Roman Empire. By recording how many Latin inscriptions report the occupation of an

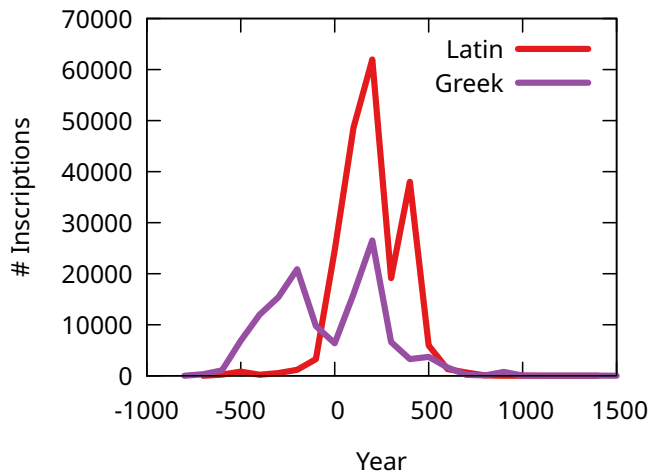


Fig. 6 | The number of inscriptions found in a given century. The number of inscriptions is on the y axis, the temporal information is on the x axis. The color reports the language: red for Latin and purple for Greek.

individual in a given Roman province, we are able to estimate the economic complexity of the areas of the empire covered by our dataset. After correcting for a number of biases, we are able to correlate the ancient economic complexity with the one estimated in modern times by using international trade data.

The main conclusion from our preliminary results is that economic complexity persists over long historic time spans. The complexity as inferred from inscriptions dating mostly to the 1st–4th centuries CE significantly correlates with the one estimated with 1962–2022 trade data (as we see from Table 2).

We should be aware that our results contradict the expectation from historiography. Before our study, we would expect a *negative* correlation between ancient and modern ECI. The countries covered by the former Eastern provinces—which were more uniformly wealthy and urbanized and therefore expected to have higher ancient complexity—currently have a lower ECI than the countries covered by the former Western provinces. However, we observe a *positive* correlation instead. There are mainly two ways to reconcile this dissonance. The first way is acknowledging that our study is biased and incomplete—due to data limitations it mostly reflects the economic complexity of the ancient Western Roman Empire—we deal with this option later in the discussion. However, we stand by our results because we have uncovered some hints that the historiography assumption that wealth and urbanization imply complexity might be wrong. We tested a correlation between ancient ECI and urbanization level, finding essentially no relation. Therefore, the second way to reconcile this dissonance is to hypothesize that, at least in ancient times, provincial urbanization and wealth do not imply complexity. An important exception to this is the capital of Rome in the province of Latium, which had an exceptionally high complexity, wealth and degree of urbanization.

Our study is correlational in nature, therefore, we cannot explain the reasons for this historical resilience in economic complexity. However, by finding evidence for it, we have now contributed to an interesting puzzle: the perseverance of the Roman economy, which was already pointed out in relation to the Roman road infrastructure^{22–24}. Explaining this resilience can open new research pathways, which can be useful to understand how we increase economic complexity today—potentially narrowing down the knowledge and economic gaps, and support developing countries.

While interesting, there are a number of factors that need to be taken into account to contextualize our results. First and foremost, when dealing with archeological and historical data, we need to be aware we are looking at an extremely biased sample. We can only analyze the inscriptions that survived up until now, which is the definition of survivorship bias. Moreover, many inscriptions that could have been produced, were not, because

the authors themselves were biased—against the epigraphic representation of women, lower-status individuals, and so on.

Crucially, even if the inscriptions were produced and survived, still more biases creep in. In order to be included in our dataset, some researchers must have worked on them, and researchers have biases as well—focusing more or less on different areas, cultures, time periods, or simply lacking the financial resources to process all the information, or develop suitable methods. We can see this issue directly in our work: we have the coded occupations for Latin inscriptions, but not for the Greek ones.

We have detailed in the paper our strategies to deal with these biases—showing how ECI rankings are robust to our best guesses on the upper and lower bounds of the occupation counts for provinces. However, our results will for sure become more accurate and robust with additional research. The most natural first step would be to properly codify the occupations from the Greek inscriptions so that we do not have to guess their content by looking at the Latin ones. Doing so, we will be able to estimate ECI also for the provinces we had to drop from our study, allowing to evaluate the extent to which our findings reflect the Roman Empire as a whole or mostly the Western Roman Empire.

There are a number of future research paths opened by this paper. The most important of them is the attempt to explain the reasons for the preservation of economic complexity over the ages. Is this because of exogenous factors—e.g., the location of a country determines its economic structure, as Italy will always be a crucial trade node in the Mediterranean—, or endogenous—i.e., being already complex gives a strong head-start to develop even more complexity in the future? Other interdisciplinary future work will involve economists: can we use what we learned about the difficulties in changing the economic landscape to foster economic growth and narrow down income inequalities across countries?

Methods

Data

For the main results of the paper, we rely on the *Latin Inscriptions in Space and Time* (L_{ALL}) dataset^{26,47}, aggregated from the major epigraphic databases of Latin inscriptions *epigraphic Database Heidelberg* and *Epigraphic Database Clauss-Slaby*. This dataset includes 525,870 Roman inscriptions in Latin, with their geographic coordinates, dates and type—e.g., epitaph, votive inscription, honorific inscription, etc., along with other attributes. The process leading to the creation of the dataset and the challenges of using the epigraphic data are discussed elsewhere^{25,27}.

In the Supplementary Materials, we test the robustness of our result by also leveraging the information from another digital dataset: *Greek Inscriptions in Space and Time* (G_{ALL})^{35,48}. This is a smaller dataset, including 217,863 inscriptions in Greek, with their geographic coordinates and dates. The original dataset provides less structured data compared to the Latin one, despite efforts to improve some of them with computer-assisted methods^{49,50}.

There are two key differences between L_{ALL} and G_{ALL} that lead us to use only L_{ALL} for the main results and G_{ALL} only for robustness. First, the two datasets span different time periods and have a different temporal spread—as Fig. 6 shows. Only about half of the Greek dataset dates to the Roman Empire, with the other half coming from the 6th to 1st centuries BCE. Merging the two datasets therefore is not straightforward and requires caution.

Moreover, L_{ALL} is structurally richer, as it has received more research attention and curation. Specifically, the L_{ALL} data set has been pre-processed by refs.^{25,27,51} to obtain the occupations (L_{OCC}), epitaphs (L_{EPI}), and people sub-datasets (L_{PEO})—Supplementary Material Section 1 provides more details about these subsets.

This is not true for the G_{ALL} dataset. Crucially for our work, we cannot reliably extract a G_{OCC} dataset comparable to L_{OCC} . The distant reading methods that would allow quantified analysis of Greek inscriptions are tailored to literary Greek language rather than the distinct epigraphic language and thus provide outputs of low quality for our purpose. Specifically, the implementation of the latest AgiLe Lemmatizer designed for Greek

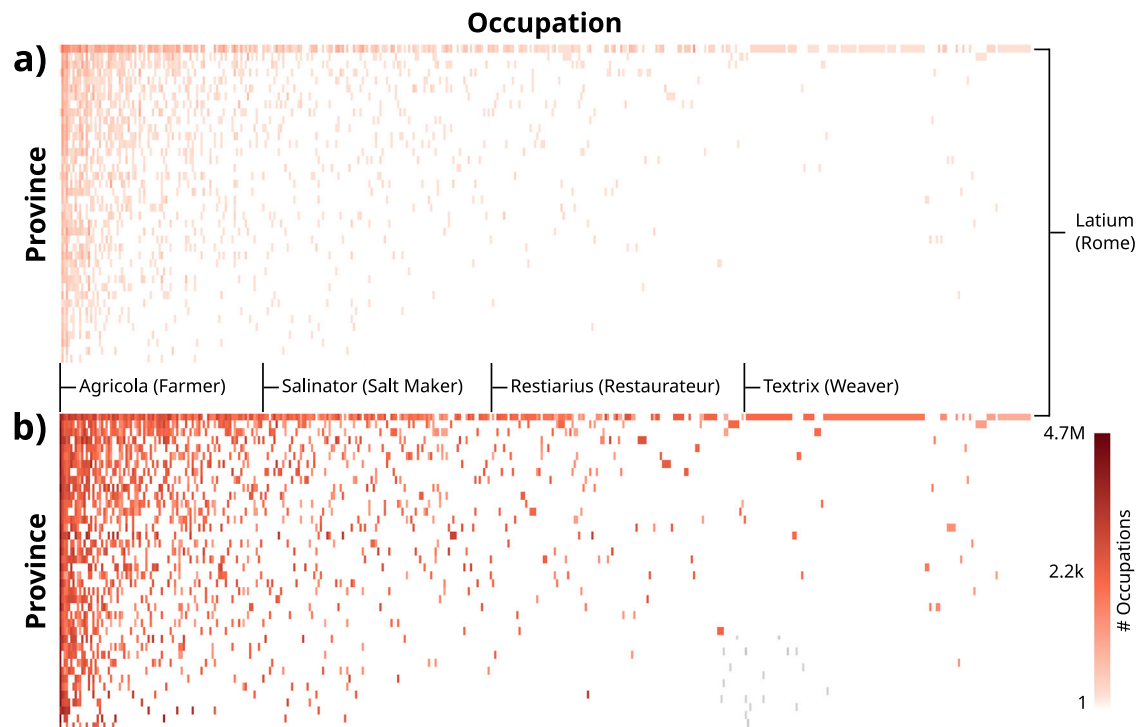


Fig. 7 | Comparing the occupational structure before and after the bias correction. **a** Reports the occupational structure before the bias correction and **b** after. Same legend as Fig. 2.

epigraphic data⁵² showed that more than 50% of occupations were detected in inscriptions where no occupation was mentioned. The results were verified by close reading of a sample of inscriptions by a human domain specialist in epigraphy (one of the authors of this paper, Heráňmanková). Until the computer-assisted detection and processing methods for Greek are improved, we have to focus on Latin as the main dataset and use Greek to inform the analysis. We detail in Supplementary Material Section 1 how we address this issue in our robustness analysis.

Other datasets used to mitigate the bias in inscription-based data include the Roman cities locations and populations, assembled by Hanson^{53,54}, and the population in 10 macroregions of the Empire (Table 3.1 in ref. 32), manually associated with the provinces that constitute them.

To calculate the modern ECI values, we rely on the UN Comtrade dataset as made available in the Harvard Dataverse⁴³. We use the *International Trade Data* (SITC, Rev 2) dataset because we want to have as wide a time interval as possible. Shorter intervals might be dominated by short-term fluctuations in global trade, which would fail to be representative of the long-term economic structures we can capture over the course of history. By calculating the ECI with data spanning decades, we are more likely to have a representative estimate.

The SITC Rev 2 data spans from 1962 to 2022, while the HS data only starts in 1995 and is therefore discarded. We calculate the ECI for this data with the exact same methodology as the one used for the ancient ECI, which is detailed later on in this section.

To calculate the modern ECI, we select a subset of countries and products. We remove all country codes not referring to actual countries, countries with less than 300k inhabitants, lower than 300 million USD GDP, or that represent less than 0.05% of the trade volume (combined). At the same time, we drop all products which taken together represent less than 2.5% of the world's trade—in practice, we preserve 97.5% of the total trade volume. This is done because the ECI calculation can become sensitive to extremely small countries and products.

Bias mitigation

Working with archeological and historical data always implies dealing with a number of biases in the data. For this work, we identified five main sources of bias. We briefly outline our strategies to deal with them here and provide further details in the Supplementary Materials Section 1.

In general, we identify reasonable bounds for our corrections, based on the current scholarship, and then we draw random samples from these distributions, leading to many different plausible M_{po} matrices.

Status

Prestigious professions and occupations elevating social status tend to be over-represented in inscriptions. At the same time, the low-prestige jobs performed mostly by slaves and household servants remain unrepresented in epigraphic data proportional to the expected demographic composition of the society^{55,56}. To counteract status bias, we inflate the counts of non-elite occupations. For the inflation factor, we draw from different distributions assuming an elite prevalence between 1% and 5% of the population, as suggested in ref. 57.

Gender

Men are disproportionately mentioned in inscriptions compared to women, regardless of their status, with roughly two men mentioned per woman. This gender unbalance is well documented in research^{30,55,56,58–62}. For this reason, we assume we need to double the number of women in our data, drawing from a distribution bounded by our best estimate about what the minimum and maximum gender disparity should be.

Primary sector

Agricultural (and other primary sector) work is absent in the inscriptions, despite it occupying roughly 85% of the population^{58,63}. We then inflate the count of the generic profession of farmer, with a proportion roughly similar to that of the rural population.

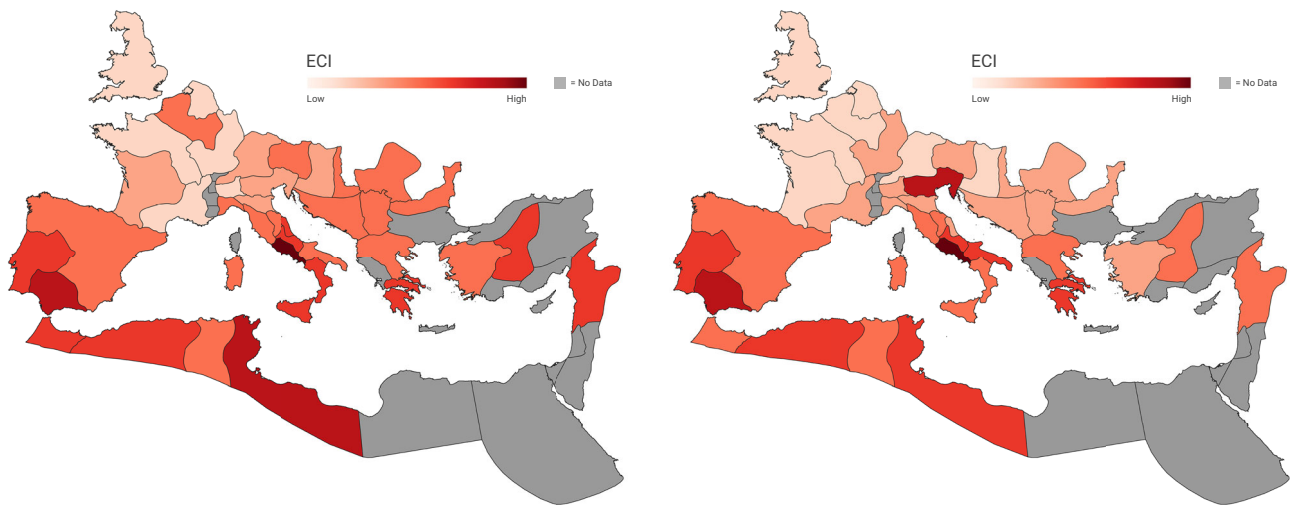


Fig. 8 | Comparing the ECI values before and after the bias correction. On the left we have the ECI estimates before bias correction and on the right after. Same legend as Fig. 3. Figure generated with QGIS.

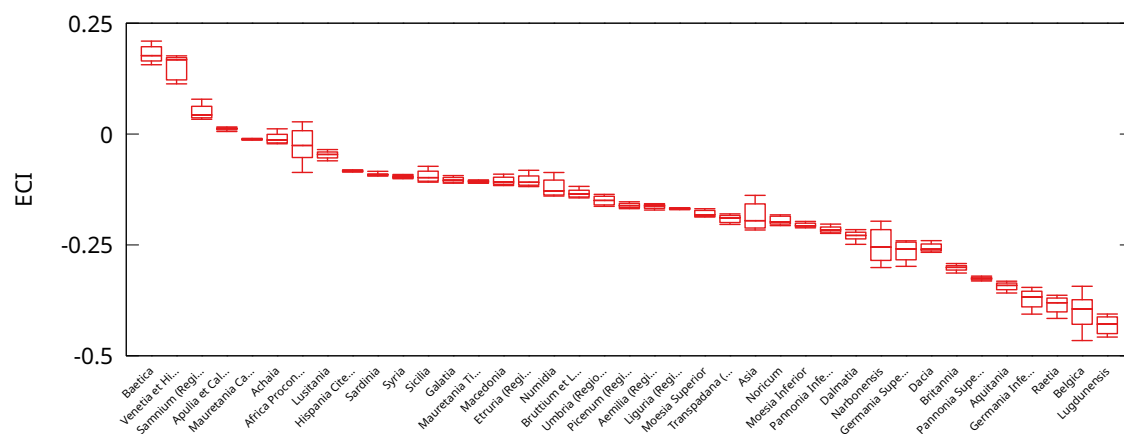


Fig. 9 | The boxplots of ECI value distributions after the bias correction per province. The boxplots show the 10th, 25th, 50th, 75th and 90th percentiles of the ECI values on the y axis. Provinces are sorted on the x axis left to right in descending order of their median ECI value. Latium is excluded.

Research

The number of inscriptions per region is not proportionate to their historical population, reflecting not only cultural differences but also the state of investigation within these regions^{64–67}. We attempt to mitigate this issue by normalizing the occupation counts to reflect the population size of each province.

The final source of bias only applies to the robustness tests in the Supplementary Material:

Language bias

The data set only contains inscriptions in Latin, but Greek was an important language of epigraphic publication, especially in the Eastern Roman Empire and some parts of the Western Roman Empire⁶⁸. Extracting occupation-related data from Greek inscriptions has not yet been implemented for technical reasons. Thus, we need to estimate the occupational data contained in them by inflating the occupation counts from the Latin inscriptions with the number of found Greek inscriptions.

Figures 7 and 8 compare the results of our pipeline without the bias correction and after the bias correction. For Fig. 7, the top half shows smaller occupation counts and white bands corresponding to occupations that are not represented in the inscriptions, such as farmers. These are included in the bias corrected version—bottom half of Fig. 7, which corresponds to Fig. 2.

For Fig. 8, we can see there is a relative shift of complexity when we correct for biases—from left (biased) to right (bias-corrected, the version we use for our main results and corresponds to Fig. 3). The bias correction leads to decreases of complexity in Gaul and the Danube limes.

Individual bias corrections have varying effects on the results, but their overall impact is comparable, except for gender corrections, which were found to have a negligible effect, as detailed in the Supplementary Materials.

Since we draw multiple potential M_{po} matrices, one question is how stable the ECI scores are across this set of possible province-occupation connections. Fig. 9 shows the distributions of ECI scores for all provinces—excluding Latium, given that it is an outlier with an extremely high ECI score which does not vary much. We can see that some provinces have a higher variability than others. However, in general, the relative rankings are reasonably stable.

Building the occupation space

The Occupation Space is a graph $G = (V, E)$. V is the set of all 440 occupations in our data—the nodes of the network. The set of edges E should contain the most significant relationships between occupations in V . Here, significance means that the two occupations co-appear in the same provinces more than one would expect at random.

To calculate the co-appearances, we first multiply M_{po} with its own transpose: $M_{oo} = M_{po}^T M_{po}$. M_{oo} gives us the raw co-appearance strength.

However, in M_{oo} strong connections could be due to chance: farmers and carpenters will have a strong connection only because they are very common occupations, but the significance of their relation might be lower than two rarer occupations that always co-occur together.

To fix this issue, we decide to use the noise-corrected backboning technique⁶⁹, which creates a null expectation for each pair of occupations in M_{oo} . We chose this technique because its null model is adequate for our data type—being developed for discrete count edge weights, which are distributed with a high skew. We then keep in E only those connections that surpass the null expectation by a significant margin.

However, this creates two issues—both of them are caused by the fact that there are several occupations that are present only in Latium. Some occupations end up not having a single connection in E . We fix this issue by picking their most significant connection and including it in E . The second issue is that the Occupation Space contains more than a single connected component. One requirement to estimate the province similarity on this network is that it must have a single connected component. Therefore, we include in E also the minimum number of the most significant edges necessary to reconnect together all connected components of the network.

In both cases, we only include edges that are stronger than the null expectation.

Economic complexity index

We calculate ECI as defined in the standard approach⁶ of the original definition³, rather than recent evolutions¹ and alternative definitions like economic fitness⁴. Economic fitness is ill-suited for data that have strong outliers like Latium, which ends up being the only province with a non-zero complexity in that framework.

For modern economies, the intuition behind ECI is that a more complex economy is able to export more than the world's average in a wide range of products. The idea is that this shows that the country possesses internally the knowledge required to create many diverse products—because diverse products like pharmaceuticals and electronic circuitry probably require different knowledge to be produced efficiently. Here “knowledge” should be intended in a wide sense: it is not only what people collectively know, but also their ability to put what they know to work—availability of capital, regulations, tacit as well as explicit knowledge, and so on^{70, 71}.

In summary, for modern economies, M_{po} is a matrix connecting each country to the products they successfully export.

This is different from our study of the economy of the Roman Empire, where M_{po} connects provinces to the occupations they employ. The difference between provinces and countries is not salient, as in modern ECI studies, there are plenty of examples using sub-national regions within the US, China, Mexico, and Australia, to name a few^{72–75}.

We argue that occupations in imperial Rome are the closest analog to products in modern economies, and that is the reason why we record occupations in our M_{po} matrix. Being a measure of knowledge, it is more suitable to use an occupation as a unit of analysis, because the tacit knowledge to perform some work is embedded in the person doing it.

However, we cannot use occupations in modern economies and have to use the proxy of the exported product, because of the fundamental difference between modern service-based economies and ancient ones. The industrial revolution shifted production from artisanal to machine-assisted, meaning that the same occupation (factory worker) could produce a wide variety of different products depending on the machine used⁷⁶. Subsequently, the transition to a service-based economy⁷⁷ (started in the 20th century and still undergoing) meant that there is a plethora of occupations that do not result in tangible goods.

Before such transitions, in pre-industrial times, the different artisans would hone their skills and specialize in producing a specific product, with little to no machine assistance. It is therefore accurate to use occupations to estimate economic complexity in pre-industrial settings, such as the Roman Empire.

The first step is to calculate the Revealed Comparative Advantage (RCA)⁷⁸ of each entry in M_{po} :

$$M_{po} = \frac{M_{po} / \sum_p M_{po}}{\sum_o M_{po} / \sum_{p,o} M_{po}}. \quad (1)$$

In practice, we compare the relative importance of occupation o for province p over the average relative importance of o across all provinces. This index is equal to 1 if o is as important to p as it is to the average of the empire. Values below 1 indicate lower than average importance, and higher than 1 a higher than average importance.

Then we binarize this transformed M_{po} by equating to 1 all $M_{po} > 1$ entries and setting zero to all other entries. In practice, we simply record the list of all o s for a province p that are present more in p than in the average Roman province.

We then produce two normalizations of this binary M_{po} : $P_{po} = M_{po} / \sum_p M_{po}$ and $O_{po} = M_{po} / \sum_o M_{po}$, which are versions of this matrix normalized by either row or column sum. The final preparation step is producing a province-province matrix:

$$M_{pp} = O_{po}^T P_{po}. \quad (2)$$

M_{pp} records the number of shared occupations between two provinces, normalized by two factors: how important they are for the two provinces and how common those occupations are in the empire at large.

Now the second largest eigenvector of M_{pp} is the ECI measure, as it is the vector capturing the largest portion of variance in the matrix. One way to interpret ECI is that it identifies the average province, the demarcation line between complex and simple provinces, and the vector measures how far a given province is from this border between two groups⁷⁹. Higher values mean farther on the side of complexity, while lower values indicate that the province is farther in the direction of simplicity.

The eigenvector does not have interpretable units, so it can be normalized in different ways—as we do in the paper—as long as the normalization is linear and does not change the ranking.

By changing the order of operations, we can also create an M_{oo} matrix, recording the normalized number of shared provinces between two occupations:

$$M_{oo} = P_{po}^T O_{po}. \quad (3)$$

By looking at the second largest eigenvector of M_{oo} we can obtain a measure of occupation complexity, just like in the M_{pp} case for provinces. We use this occupation complexity measure for Fig. 4.

Data availability

All data is available. The inscription data is available on Zenodo at <https://zenodo.org/records/10473706> and <https://doi.org/10.5281/zenodo.10139110>. The modern trade data is available in the Harvard Dataverse at 10.7910/DVN/H8SFD2&version=10.0 and 10.7910/DVN/3BAL10&version=6.0.

Code availability

The code is available on our GitHub: https://github.com/past-networks/roman_ECI/.

Received: 26 September 2025; Accepted: 21 September 2026;

Published online: 06 October 2026

References

- Hidalgo, C. A. Economic complexity theory and applications. *Nat. Rev. Phys.* **3**, 92–113 (2021).
- Balland, P.-A. et al. The new paradigm of economic complexity. *Res. Policy* **51**, 104450 (2022).
- Hidalgo, C. A. & Hausmann, R. The building blocks of economic complexity. *Proc. Natl. Acad. Sci. USA* **106**, 10570–10575 (2009).

4. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A. & Pietronero, L. A new metrics for countries' fitness and products' complexity. *Sci. Rep.* **2**, 723 (2012).
5. Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G. & Pietronero, L. Measuring the intangibles: a metrics for the economic complexity of countries and products. *PloS One* **8**, e70726 (2013).
6. Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M. & Simoes, A. *The Atlas of Economic Complexity: Mapping Paths to Prosperity* (MIT Press, 2014).
7. Hidalgo, C. A. et al. The principle of relatedness. In *Unifying Themes in Complex Systems IX: Proceedings of the Ninth International Conference on Complex Systems* 9, 451–457 (Springer, 2018).
8. Coscia, M., Neffke, F. M. & Hausmann, R. Knowledge diffusion in the network of international business travel. *Nat. Hum. Behav.* **4**, 1011–1020 (2020).
9. Pinheiro, F. L., Hartmann, D., Boschma, R. & Hidalgo, C. A. The time and frequency of unrelated diversification. *Res. Policy* **51**, 104323 (2022).
10. Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristrán, M. & Hidalgo, C. A. Linking economic complexity, institutions, and income inequality. *World Dev.* **93**, 75–93 (2017).
11. Mealy, P. & Teytelboym, A. Economic complexity and the green economy. *Res. Policy* **51**, 103948 (2022).
12. Hidalgo, C. A. The policy implications of economic complexity. *Res. Policy* **52**, 104863 (2023).
13. Zhu, S. & Li, R. Economic complexity, human capital and economic growth: empirical research based on cross-country panel data. *Appl. Econ.* **49**, 3815–3828 (2017).
14. Brughmans, T. Thinking through networks: a review of formal network methods in archaeology. *J. Archaeol. Method Theory* **20**, 623–662 (2013).
15. Knappett, C. *Network Analysis in Archaeology: New Approaches to Regional Interaction* (Oxford University Press, 2013).
16. Collar, A., Coward, F., Brughmans, T. & Mills, B. J. Networks in archaeology: phenomena, abstraction, representation. *J. Archaeol. Method Theory* **22**, 1–32 (2015).
17. Peeples, M. A. Finding a place for networks in archaeology. *J. Archaeol. Res.* **27**, 451–499 (2019).
18. Bowman, A. & Wilson, A. Quantifying the Roman economy. In *Quantifying the Roman Economy: Methods and Problems*, Oxford Studies in the Roman Economy, 3–84 (Oxford University Press, Oxford, New York, 2009).
19. Brughmans, T. & Wilson, A. (eds.) *Simulating Roman Economies: Theories, Methods, and Computational Models* (Oxford University Press, Oxford, New York, 2022).
20. Ortman, S. G. et al. Identification and measurement of intensive economic growth in a Roman imperial province. *Sci. Adv.* **10**, eadk5517 <https://doi.org/10.1126/sciadv.adk5517> (2024).
21. Flückiger, M., Hornung, E., Larch, M., Ludwig, M. & Mees, A. Roman transport network connectivity and economic integration. *Rev. Econ. Stud.* **89**, 774–810 <https://doi.org/10.1093/restud/rdab036> (2022).
22. Wahl, F. Does European development have Roman roots? Evidence from the German Limes. *J. Econ. Growth* **22**, 313–349, <https://doi.org/10.1007/s10887-017-9144-0> (2017).
23. Dalgaard, C.-J., Kaarsen, N., Olsson, O. & Selaya, P. Roman roads to prosperity: persistence and non-persistence of public infrastructure. *J. Comp. Econ.* **50**, 896–916 <https://doi.org/10.1016/j.jce.2022.05.003> (2022).
24. De Benedictis, L., Licio, V. & Pinna, A. M. From the historical Roman road network to modern infrastructure in Italy. *J. Regional Sci.* **63**, 1162–1191 <https://doi.org/10.1111/jors.12659> (2023).
25. Kaše, V., Heráňková, P. & Sobotková, A. Division of labor, specialization and diversity in the ancient Roman cities: a quantitative approach to Latin epigraphy. *PloS One* **17**, e0269869 (2022).
26. Kaše, V., Heráňková, P. & Sobotková, A. List, <https://doi.org/10.5281/zenodo.8431323> (2023).
27. Heráňková, P., Kaše, V. & Sobotková, A. Inscriptions as data: digital epigraphy in macro-historical perspective. *J. Digital Hist.* **1**, 99–141 <https://doi.org/10.1515/jdh-2021-1004> (2021).
28. Tabash, M. I., Mesagan, E. P. & Farooq, U. Dynamic linkage between natural resources, economic complexity, and economic growth: empirical evidence from Africa. *Resour. Policy* **78**, 102865 (2022).
29. Pinheiro, F. L., Balland, P.-A., Boschma, R. & Hartmann, D. The dark side of the geography of innovation: relatedness, complexity and regional inequality in Europe. *Regional Stud.* **59**, 2106362 (2025).
30. Saller, R. Household and Gender. In Morris, I., Saller, R. P. & Scheidel, W. (eds.) *The Cambridge Economic History of the Greco-Roman World*, 87–112 <https://doi.org/10.1017/CHOL9780521780537.005> (Cambridge University Press, Cambridge, 2007).
31. Salomies, O. *Epigraphic Evidence: Ancient History from Inscriptions*, chap. Names and Identities: Onomastics and Prosopography (Routledge, 2001).
32. Scheidel, W. Demography. In Scheidel, W., Morris, I., Saller, R. P. et al. (eds.) *The Cambridge Economic History of the Greco-Roman World*, vol. 1, 38–86 (Cambridge University Press Cambridge, 2007).
33. Sobotková, A., Heráňková, P. & Kaše, V. Soldier mobility in the Roman Empire on the basis of origo attestations in Latin epigraphic evidence. in *Roads, Economy, and Travel in the Roman World* (Brughmans, T. & Pažout, A. eds) (tba, n.d.).
34. Centre, A. W. M. Provinces of the roman empire in ce 200 (2016).
35. Kaše, V., Heráňková, P. & Sobotková, A. Gist, <https://doi.org/10.5281/zenodo.10139110> (2023).
36. Patterson, B. D. & Atmar, W. Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biol. J. Linn. Soc.* **28**, 65–82 (1986).
37. Atmar, W. & Patterson, B. D. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* **96**, 373–382 (1993).
38. Hausmann, R. & Hidalgo, C. A. The network structure of economic output. *J. Econ. Growth* **16**, 309–342 (2011).
39. Bustos, S., Gomez, C., Hausmann, R. & Hidalgo, C. A. The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS One* **7**, e49393 (2012).
40. Neffke, F. & Henning, M. Skill relatedness and firm diversification. *Strategic Manag. J.* **34**, 297–316 (2013).
41. Van Leeuwen, M. et al. *HISCO. Historical International Standard Classification of Occupations* (Leuven University Press, 2002).
42. Leeuwen, V. Files from HISCO database, 10622/JA9B80 (2016).
43. at Harvard University, T. G. L. *International Trade Data* (SITC, Rev. 2), <https://doi.org/10.7910/DVN/H8SFD2> (2019).
44. Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317**, 482–487 (2007).
45. Coscia, M. Generalized Euclidean measure to estimate network distances. In *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 119–129 (2020).
46. Damstrup, A. S. R., Madsen, S. T. & Coscia, M. Unsupervised learning via network-aware embeddings. *arXiv preprint arXiv:2309.10408* (2023).
47. Kaše, V., Heráňková, P. & Sobotková, A. List, <https://doi.org/10.5281/zenodo.10473706> (2024).
48. Kaše, V. & Heráňková, P. Gist, <https://doi.org/10.5281/zenodo.6827729> (2022).
49. Sommerschild, T. et al. I.PHI dataset: ancient Greek inscriptions (2021).
50. Assael, Y. et al. Restoring and attributing ancient texts using deep neural networks. *Nature* **603**, 280–283 <https://doi.org/10.1038/s41586-022-04448-z> (2022).

51. Kaše, V., Heráňmanková, P. & Sobotkova, A. Classifying Latin Inscriptions of the Roman Empire: a machine-learning approach. *Proc. 2nd Workshop Comput. Humanit. Res. (CHR2021)* **2989**, 123–135 (2021).
52. de Graaf, E., Stopponi, S., Bos, J. K., Peels-Matthey, S. & Nissim, M. AGILe: The First Lemmatizer for Ancient Greek Inscriptions. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (Calzolari, N. et al. eds.) 5334–5344 (European Language Resources Association, Marseille, France, 2022).
53. Hanson, J. W., Ortman, S. G., Bettencourt, L. M. & Mazur, L. C. Urban form, infrastructure and spatial organisation in the Roman Empire. *Antiquity* **93**, 702–718 <https://doi.org/10.15184/aqy.2018.192> (2019).
54. Hanson, J. W. & Ortman, S. G. Population, area, and infrastructural measures for roman cities of the imperial period. tDAR id: 448563, <https://doi.org/10.6067/XCV8448563> (2019).
55. Joshel, S. R. *Work, Identity and the Legal Status at Rome: Study of the Occupational Inscriptions* (University of Oklahoma Press, Oklahoma, 1993).
56. Treggiari, S. Jobs in the Household of Livia. *Pap. Br. Sch. Rome* **43**, 48–77 (1975).
57. Scheidel, W. & Friesen, S. J. The size of the economy and the distribution of income in the Roman Empire. *J. Rom. Stud.* **99**, 61–91 (2009).
58. Scheidel, W. The most silent women of Greece and Rome: rural labour and women's life in the ancient world (I). *Greece Rome* **42**, 202–217 (1995).
59. Scheidel, W. The Most Silent Women of Greece and Rome: Rural Labour and Women's Life in The Ancient World (II). *Greece Rome* **43**, 1–10, <https://doi.org/10.1093/gr/43.1.1> (1996).
60. Boserup, E. *Woman's Role in Economic Development* (St. Martin's Press, New York, 1970).
61. Becker, H. Roman women in the urban economy: occupations, social connections, and gendered exclusions Edited by J. Turfa and S. Budin. London: Routledge,. In Turfa, J. & Budin, S. (eds.) *Women in Antiquity: Real Women Across the Ancient World*, 915–931 (Routledge, London, 2016).
62. Treggiari, S. Lower class women in the Roman economy. *Florilegium* 65–86 (1979).
63. Bowman, A. & Wilson, A. (eds.) *The Roman Agricultural Economy: Organisation, Investment, and Production* (Oxford University Press, 2013).
64. Bodel, J. Epigraphic Culture and the Epigraphic Mode. In Benefiel, R. & Keesling, C. M. (eds.) *Inscriptions and the Epigraphic Habit*, 1–44, https://doi.org/10.1163/9789004683129_002 (Brill, Leiden; Boston, 2024).
65. Mullen, A. & Willi, A. (eds.) *Latinization, Local Languages, and Literacies in the Roman West* (Oxford University Press, Oxford, 2024).
66. Woolf, G. Monumental writing and the expansion of Roman society in the early empire. *J. Rom. Stud.* **86**, 22–39 <https://doi.org/10.2307/300421> (1996).
67. Meyer, E. A. Epigraphy and Communication. In *The Oxford Handbook of Social Relations in the Roman World* (Peachin, M. ed)(Oxford University Press, Oxford, 2011).
68. Prag, J. R. Epigraphy in the western Mediterranean: a Hellenistic phenomenon? In *The Hellenistic West* (Prag, J. R. W. & Quinn, J. C. eds.) 320–347, <https://doi.org/10.1017/CBO9781139505987.012> (Cambridge University Press, 2013).
69. Coscia, M. & Neffke, F. M. Network backboning with noisy data. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, 425–436 (IEEE, 2017).
70. Polanyi, M. *Personal Knowledge* (University of Chicago Press, Chicago, IL, 1974).
71. Nonaka, I. & Takeuchi, H. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics Of Innovation* (Oxford University Press, Oxford, England, 1995).
72. Chávez, J. C., Mosqueda, M. T. & Gómez-Zaldívar, M. Economic complexity and regional growth performance: evidence from the Mexican economy. *Rev. Regional Stud.* **47**, 201–219 (2017).
73. Reynolds, C. et al. A sub-national economic complexity analysis of Australia's states and territories. *Regional Stud.* **52**, 715–726 (2018).
74. Gao, J. & Zhou, T. Quantifying China's regional economic complexity. *Phys. A Stat. Mech. Appl.* **492**, 1591–1603 (2018).
75. Fritz, B. S. & Manduca, R. A. The economic complexity of us metropolitan areas. *Regional Stud.* **55**, 1299–1310 (2021).
76. Landes, D. S. *The Unbound Prometheus* (Cambridge University Press, 1969).
77. Wren, A. *The Political Economy of the Service Transition* (Oxford University Press (UK), 2013).
78. Balassa, B. Trade liberalisation and “revealed” comparative advantage 1. *Manch. Sch.* **33**, 99–123 (1965).
79. Mealy, P., Farmer, J. D. & Teytelboym, A. Interpreting economic complexity. *Sci. Adv.* **5**, eaau1705 (2019).

Acknowledgements

The shapefile of Roman provinces in 200 CE (peak Roman Empire under the Severan dynasty) is based on 'roman-empire-ce-200-provinces.geojson' published by The Ancient World Mapping Center and corrected by Adam Pažout in 2023. This research was conducted within the framework of the Past Social Networks Project (2023–2026), funded by The Carlsberg Foundation's Young Researcher Fellowship (CF21-0382). This work was supported by a research grant (VIL57402) from VILLUM FONDEN.

Author contributions

M.C. and T.B. designed the study. P.H. collected the data. M.M., P.H. and M.C. wrote code and prepared the figures. All authors wrote and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44260-026-00100-3>.

Correspondence and requests for materials should be addressed to Michele Coscia.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.