Cities, information, and the epigraphic habit: re-evaluating the links between the numbers of inscriptions and the sizes of sites

J. W. HANSON

Cities, Information, and the Epigraphic Habit

Re-evaluating the Links between the Numbers of Inscriptions and the Sizes of Sites

ABSTRACT Among classical scholars there is a widespread assumption that there is no relationship between the sizes of communities and their epigraphic output. In this article, I offer a new model, which suggests two hypotheses for how inscriptions increase with population, depending on whether they can be regarded as a form of infrastructure or a measure of wealth or disposable income. I show that, despite the variation between sites, there is nonetheless a consistent relationship between the numbers of inscriptions and the estimated populations of sites. The numbers of inscriptions increase slower than the estimated populations of sites, however, suggesting that they acted as a form of information infrastructure. This has important implications for our understanding of the mechanisms for transmitting information in ancient contexts, suggesting several avenues for future research.

KEYWORDS Complex systems; settlement scaling theory; inscriptions; epigraphic culture; urbanism; Roman Empire

Acknowledgements and Funding Details

I would like to thank Scott Ortman for discussing ideas and Annalisa Marzano, Jose Lobo, and the anonymous reviewers for reading and commenting on earlier drafts. This research was funded by a British Academy Postdoctoral Fellowship (PF2/180110).

Introduction

Over the last decade, scholars have increasingly highlighted the importance of regarding cities as complex systems, emphasizing the roles played by flows of material, energy, and information. Having said this, although it would be natural to assume that there is a close relationship between the sizes of Greek and Roman cities and the numbers of inscriptions documented in urban contexts, this attitude has not, in fact, been the dominant one among classical archaeologists and ancient historians, who have generally argued that there is very little relationship between the sizes of communities and their epigraphic outputs (Duncan-Jones 1982; MacMullen 1982; Beltrán Lloris 2014). This scepticism is based on two components. The first is that there were significant differences in the extent to which individuals chose to use inscriptions across space and time.[1] As Ramsay MacMullen (1982, 233) explained in an important article nearly forty years ago, even at the peak of the Roman Empire,

> we must look not for the occasion chosen, such as a life ended, a vow made, or an honor voted, but at the decision itself to give those facts some […] commemoration […] even in the Roman Empire, there were deaths, vows, and decrees unrecorded, more or less often in different times and places.

Although this observation is normally made in the context of funeral inscriptions, there is no reason why it should not also apply to the whole range of inscriptions. These differences are the result of variation not only in wealth and status, but also literacy rates, custom, and fashion, as well as discrepancies in the amount of material that has been preserved, recovered, and published (Duncan-Jones 1982; MacMullen 1982; Beltrán Lloris 2014). This has inevitably led to doubts about the extent to which the occurrence of inscriptions can be seen as a reliable reflection of wider social and economic conditions and whether

1 For general discussions about these issues, see MacMullen 1982; Saller and Shaw 1984; Mann 1985; Meyer 1990; Cherry 1995; Woolf 1996; Bodel 2001; Keppie 1991; Harvey 2004; Cooley 2012; Bruun and Edmondson 2015.

Dr J. W. Hanson (j.w.hanson@reading.ac.uk) is a British Academy Postdoctoral Fellow in the Department of Classics at the University of Reading, UK. ORCID iD: 0000-0003-4480-4791.

 Journal of Urban Archaeology, 4 (2021), pp. 137–152
BREPOLS PUBLISHERS
10.1484/J.JUA.5.126597

they can be regarded as a proxy for socially important information or wealth. The second component of this scepticism is that there are significant differences in the survival rates of inscriptions themselves (Duncan-Jones 1982; Beltrán Lloris 2014). This, in itself, is influenced by various factors, including not only the amount of each settlement that has been excavated, but also the durability of the materials used, the extent to which different materials were reused, how far continued occupation has erased earlier remains, and whether specific areas have been unearthed (given that some areas yield more inscriptions than others) (Duncan-Jones 1982, 361). Most scholars have therefore concluded that it is very unlikely that there is any relationship between the sizes of sites and the numbers of inscriptions associated with them. As Richard Duncan-Jones (1982, 360) has put it, the fact that 'town A' has left more inscriptions than 'town B' almost never indicates, in itself, that A was larger or wealthier than B.

There are various reasons, however, for re-evaluating this view. There is now widespread evidence for systematic relationships between the sizes of sites and some of their most important attributes in a range of contexts, including both ancient and modern sites in both the Old and the New World, which seem to be surprisingly uniform from context to context.[2] One of the most important findings is that the inhabitants of larger cities, on average, not only live at higher population densities than those in smaller cities, but also have a larger number of social contacts, again on average (Bettencourt and West 2010; West 2017; Smith 2018; Lobo and others 2020; Ortman and others 2020). In other words, increasing densities lead to increasing interactions. This, in turn, creates increasing opportunities for individuals to exchange knowledge, skills, and ideas, enabling economies of scale with infrastructure and increasing returns to scale with various social and economic measures, including a generally faster pace of life (Bettencourt and West 2010). This resonates with earlier work in urban studies, which has suggested that the defining features of cities are social networks, which are used by their inhabitants to generate and share information (Meier 1962). This has significant implications for our understanding of urban life, since it means that one would not only expect larger cities to generate more information, but also to be home to a larger number of activities, which one would expect to be reflected in material remains, including inscriptions. Although these kinds of relationships cannot be used to make predictions about individual sites, they can be used to characterize the underlying dynamics of urban systems, so long as a wide enough sample of sites is assessed. This line of reasoning suggests that, contrary to existing beliefs, it should be possible to find evidence for a general relationship between the sizes of communities and their epigraphic cultures, even if it is not possible to predict the exact values for Duncan-Jones's 'town A' and 'town B'.

Although some of these relationships have been observed in contemporary cities, using various proxies, it is not known whether they also occurred in non-modern contexts. In addition, while there is now extensive evidence for consistent relationships between the estimated populations of ancient cities and various aspects of the built environment (e.g. the dimensions of *fora*, *agorai*, street networks, city gates, and seating capacities of theatres and amphitheatres), there is so far little evidence that these relationships can be extended to specific kinds of artefacts, such as inscriptions (which could then be seen as a reflection of social and economic conditions) (Hanson and others 2019; Hanson 2020; Hanson and Ortman 2020). The Greek and Roman world should give us an excellent opportunity to test these ideas, given the development of new estimates for the populations of sites in recent times and the large numbers of inscriptions that could be used as a proxy for information, in their role as bearers of socially important meaning. This therefore raises the questions of 1) whether the current scepticism about the links between the sizes of sites and the numbers of inscriptions found in them is justified or whether it is the result of a lack of systematic analysis and 2) what implications this has for our understanding of the mechanisms through which information was transmitted in non-modern settlements.

In this article, I will re-evaluate inscriptions and the links between the sizes of communities and their epigraphic outputs, focusing on the roles of inscriptions as sources of information. To do this, I will begin by offering a new model for how one would expect inscriptions, and the information contained in them, to increase with the sizes of cities, drawing on recent developments in complex systems, settlement scaling theory, and information studies (Bettencourt 2013; 2014; Schläpfer and others 2014; Lobo and others 2020). Having established these concepts, I will then use a combination of existing material, in the form of the *Epigraphik-Datenbank Clauss / Slaby*, and statistical techniques to evaluate the relationship between the total numbers of inscriptions and the estimated populations of the sites. This allows me to assess not only how much the empirical evidence matches theoretical expecta-

2 Lobo and others 2020. For examples, see Table 7.2.

tions, but also how quickly or slowly the numbers of inscriptions change with the sizes of sites (the slope of the relationship between them), what the baseline numbers of inscriptions per person was (the y-intercept of the relationship), how much each site deviates from the overall relationship (the residuals of the relationship), and whether these deviations have a geographical or chronological component (this is known as autocorrelation). I will show that there is in fact a relationship between the numbers of inscriptions associated with sites and their estimated populations, despite the considerable variation between individual cases. This relationship is not linear, however, and the total numbers of inscriptions associated with sites grow more slowly than their estimated populations. In other words, there was a decreasing number of inscriptions per capita as the sizes of cities increased. This indicates that inscriptions acted as a kind of information infrastructure and were not, by themselves, a reflection of increasing wealth or disposable income. This suggests that both of the perspectives discussed above are partly right, in the sense that, although there is a broad relationship between the numbers of inscriptions associated with sites and their estimated populations, which supports the notion of an overall 'epigraphic habit', there is a large amount of variance in this relationship, revealing significant differences in the epigraphic cultures of sites. These results therefore lay the groundwork for more detailed work in the future.

Theoretical Background

As recent work has shown, although it can be difficult to measure the total amount of information circulating in settlements, it is possible to use insights from recent research on complex systems to suggest a model for the relationship between the amount of information and the sizes of settlements (Hanson and Ortman 2020). This research, which is known as settlement scaling theory, has not only shown that there is a series of relationships between the sizes of settlements and some of their most important attributes, with various social and economic measures tending to increase faster than the populations of settlements (i.e. in a superlinear fashion) and different kinds of infrastructure tending to increase slower than the populations of settlements (i.e. in a sublinear fashion), but also that the slopes of these relationships tend to conform to a narrow range of values (Lobo and others 2020). These relationships have been explained by the roles of settlements as 'social reactors', which concentrate people in space and time and therefore increase the opportunities for them to interact, share resources, and exchange knowledge, skills, and ideas (Bettencourt 2013). These patterns are not just a product of the numbers of inhabitants, but also result from the increasing population densities and social interactions associated with increasing size. This has allowed scholars to suggest a series of formal mathematical models for these relationships, which are based on the central observation that settlements can be regarded as social networks embedded in the built environment (Lobo and others 2020).

These models have important implications, since they not only suggest that the amount of information in a settlement is proportional to the total number of social interactions within it, but also that the transmission of information, at a fundamental level, requires some form of infrastructure. This means that one would expect the amount of information circulating in a settlement to be a product of three variables: the number of sources of information, the number of times they were consulted, and the number of times their content was repeated from person to person. As José Lobo and others (2020) have shown, since the total number of social interactions tends to increase faster than the populations of sites, one would also expect the amount of information to increase faster than the populations of settlements. In addition, given that the residents of larger cities tend to have had a greater number of social contacts, on average, than those of smaller cities, and tend to live at higher population densities, one would expect the same source of information to have been viewed and repeated more frequently, meaning that the mechanisms for transmitting information (i.e. information infrastructure) should increase more slowly than the number of residents. This means that, although one would expect the amount of information to exhibit increasing returns to scale, one would expect the infrastructure for conveying that information to exhibit economies of scale.

These suggestions can then be refined even further, given that there are now both theoretical and empirical reasons for believing that the total number of social interactions and the average densities of settlements tend to increase at a specific rate, with an exponent of four-thirds in both cases, while infrastructure tends to increase with an exponent of two-thirds (Lobo and others 2020). This observation has been borne out by recent work on the Greek and Roman world, which shows that the population densities of settlements tend to increase faster than their populations, with an exponent of about four-thirds, while the dimensions of both mixing spaces, such as *fora* and *agorai*, and street networks tend to increase more slowly than their populations, at a rate of about two-thirds (Hanson and Ortman 2017; Hanson and others 2019). It is therefore reasonable to assume that any infrastructure for conveying information would increase more slowly than population, with an exponent of about two-thirds. This can then be turned into a simple formal model, drawing on wider settlement scaling theory, to give:

$$I = aN^{\beta}$$

where I is the infrastructure associated with information, a is a constant baseline value, N is the population of the settlement, and β is an exponent, with a value of about two-thirds.

This model might have some explanatory power for inscriptions, given that their main purpose was to convey socially important information, some of which was expressed through the setting up of the inscription itself (and any other associated structures or events) and some of which was expressed through the text of the inscription and any details it recorded. This means that one might also expect inscriptions to increase slower than the populations of sites, at the rate of two-thirds. It is important, however, to make a number of caveats. First, although one would expect the potential audience of these inscriptions to be the whole settlement, it is important to acknowledge that not everyone would have been able to read them, given the literacy rate of the ancient world. It is difficult to incorporate this into the model, given the lack of evidence for whether or how these rates changed with the sizes of cities. Secondly, although this article can be regarded as an extension of earlier studies, there are obviously important differences between the material that is discussed here and the material that has been discussed elsewhere (Hanson and Ortman 2020). The most important of these are that these inscriptions represent a series of events, rather than just one, and are associated with multiple audiences, instead of one. This means that one would not expect inscriptions to be associated with a fixed amount of information or a fixed number of sources of information, unlike theatres or amphitheatres, where each spectator saw the same event and the number of venues was limited. Thirdly, inscriptions did not just record facts about events, but also had a symbolic function. This means that they should not be viewed in purely functional terms, but were also an important social and cultural phenomenon in their own right. Fourthly, it is important to recognize that there are several different kinds of inscriptions, which were set up by different individuals for different audiences. This means that one might expect different kinds of inscriptions to scale at different rates, with more personal inscriptions being closer to linear and more general ones closer to sublinear (for example, consider the difference between epitaphs and other notices). Although it is possible, in principle, to suggest a number of models for different kinds of inscriptions, this is beyond the scope of this article, which is concerned with the overall relationship between the numbers of inscriptions and the sizes of sites. I have therefore regarded inscriptions as part of the same basic social and cultural phenomenon, allowing us to focus on the shared roles they played in how settlements functioned. Finally, it is also important to remember that inscriptions were not the only source of information. This means that, although it is legitimate to regard them as an important mechanism through which information was transmitted in ancient settlements, it would be misguided to use them as a simple proxy for the total amount of information that was contained within a settlement.

This framework also allows us to consider the possibility that the number of inscriptions is a reflection of the overall wealth of settlements. As Lobo and others (2020) have shown, since one would expect the total amount of wealth in a settlement to be proportional to the total number of social interactions within it, one would also expect the total amount of wealth that is generated by settlements to increase faster than their populations, again with an exponent of around four-thirds. This means that the total number of inscriptions would increase faster than the populations of settlements if they are simply a reflection of the total amount of wealth or disposable income in settlements.

These considerations can then be combined to suggest two competing hypotheses for how one would expect the numbers of inscriptions to increase with the estimated populations of sites. If inscriptions can be regarded as a form of infrastructure for conveying information one would expect them to increase slower than the populations of sites with an exponent of two-thirds. In contrast, if they are simply a reflection of wealth or disposable income, one would expect them to increase faster than the populations of sites with an exponent of four-thirds. These hypotheses can then be tested by exploring the relationship between the numbers of inscriptions and the estimated populations of sites and examining the slope of the relationship.

At this point, it is worth reiterating that the relationship that is discussed below is only intended to capture the average conditions across settlements and cannot necessarily be used to make predictions about individual sites, which deviate from these relationships to various degrees. As has been noted elsewhere, although these deviations might be a result of errors, such as in how the numbers of inscriptions have been counted or in how the sizes of sites have been estimated, it is also possible to regard them as a meaningful reflection of the individual social and economic conditions of sites, once the overall effects of their size have been taken into account (Hanson and Ortman 2020). In other words, although one would expect the sizes of settlements to have the greatest impact on their overall characteristics, one would also expect there to be secondary factors that govern the under- or overperformance of sites. This is illustrated, for example, by recent work on theatres and amphitheatres, which suggests that, although there is an overall relationship between the seating capacities of these structures and the estimated populations, there is also a significant amount of deviation from the overall relationship, which appears to be explained by the civic statuses of sites (provincial capitals tend to overperform relative to their size (i.e. to deviate positively), while *municipia* tend to underperform (i.e. to deviate negatively)) (Hanson and Ortman 2020). As will be seen below, this approach therefore gives us a powerful way of thinking about inscriptions, since it allows us to appreciate the effects of the sizes of sites on the numbers of inscriptions, without discounting the possibility that there is variation from site to site, which reflects different amounts of evidence or different epigraphic traditions.

Inscriptions and Estimated Populations

By any standards there is a vast amount of epigraphic material in the ancient world. Although it is impossible to come up with an exact figure, scholars have suggested a conservative estimate of at least half a million inscriptions, coming from over twenty thousand locations (Beltrán Lloris 2014, 136). Most of these are concentrated in Rome and Italy, while most date to the early imperial period (Beltrán Lloris 2014, 137). These inscriptions relate to almost all aspects of ancient life and cover a wide range of subject matters. They are therefore a crucial source for various aspects of the ancient world, including both demographic questions, such as average ages at death and levels of migration, and socio-economic ones, such as the amount that was spent on the construction, maintenance, and repair of buildings, etc. As Francisco Beltrán Lloris (2014, 136) has pointed out, however, a significant number of inscriptions are either quite short or fragmentary, which means that they are often not that revealing. In fact, most inscriptions are fewer than ten words long (Beltrán Lloris 2014, table 8.1). This issue is discussed below.

I have used the *Epigraphik-Datenbank Clauss / Slaby* (*EDCS*), since it is the most comprehensive source available at the time of writing.[3] It includes nearly all the Latin inscriptions that have been published over the last couple of centuries. These are drawn from a wide range of sources, including the *Corpus inscriptionum latinarum* (*CIL*), a monumental project started in the nineteenth century, which itself includes about 180,000 inscriptions, spread across seventeen volumes, as well as inscriptions from various regional compendia and the annual updates of *L'Année épigraphique* (Beltrán Lloris 2014, 136). Although there are other resources that could be used, such as the *Epigraphische Datenbank Heidelberg* (*EDH*), these are not as complete, since they are either still being compiled or only focus on certain regions and/or periods. As Andrew Wilson (2011, 167) has noted, the coverage of these resources is also somewhat inconsistent across different regions and/or periods. According to him, although the *EDH* is 'good for the Balkans, Achaia, Dacia, Dalmatia, Epirus, Macedonia, Moesia Inferior, Moesia Superior, Thracia' and 'fairly good for Spain', it is still 'patchy for Italy' and 'very poor for Africa, Asia, and Gaul' (Wilson 2014, 164). The *EDCS* is also available online and gives the full text of each inscription (but not translations). These texts are based on the regularized transcriptions that are given by the original sources, but also include editorial expansions of the abbreviations that were used in inscriptions, so they are not an exact reproduction of what was originally inscribed. Although some attempts have been made to categorize these inscriptions (such as into votive, funerary, or building inscriptions), many of these categories are quite broad, which means that it is more productive to search them for specific words or phrases, allowing us to focus on specific themes (this is illustrated below). Finally, the *EDCS* also records the find-spots of nearly all inscriptions and the date ranges of a significant number of them.

Although the *EDCS* is an active database, which is continually being updated and supplemented, it included a total of 503,766 individual inscriptions at the time the research for this article was undertaken, which relate to about 20,082 find-spots. These cover a wide region, including most of the Mediterranean and its periphery, and a wide period, from the third century BC to the eighth century AD. These inscriptions can be very broadly divided into a number of

3 See <http://db.edcs.eu/epigr/epi.php?s_sprache=en> [accessed 7 July 2020].

types, which include not only building or dedicatory inscriptions, but also epitaphs, honorific inscriptions, legal texts, votive inscriptions, milestones, boundary inscriptions, statements of ownership, and labels, although there are many others (the database does not include wax tablets or other forms of communication). While most inscriptions are derived from cities and towns, some are also derived from religious and military sites, including frontier installations such as Hadrian's Wall. Although nearly all inscriptions can be located, only a small fraction can be dated, since very few of their commissioners insisted on including this information in the inscription. As a result, most inscriptions are dated on the basis of their content (e.g. references to specific individuals or events) or on the basis of their context and style (e.g. carving and letter forms). The dates of most inscriptions are therefore quite approximate and are often expressed in terms of wide date ranges.

There are two issues with the *EDCS*, however, which need to be briefly considered. The first is that it also includes a significant amount of writing on everyday objects, which is traditionally referred to as *instrumentum domesticum* by classical archaeologists and ancient historians. Although it is not possible to estimate the exact amount of *instrumentum domesticum* that is included in the database, since this is often not recorded, it is possible to draw on more recent studies to get a rough sense of the proportion of this material. This shows that, although the amount of *instrumentum domesticum* that is included in the database varies significantly, the average is about 18 per cent (Beltrán Lloris 2014, table 8.2). Most of the cases with the highest percentages of *instrumentum domesticum* are in the north-western provinces, including Britannia, Germania Superior and Inferior, and Gallia Aquitania, Belgica, and Lugdunensis. The second issue is that the *EDCS* contains a significant number of duplicates (Beltrán Lloris 2014, 136). This is most marked with *instrumentum domesticum*, given that every example is included, even if it is derived from the same stamp. Interestingly, both problems seem to be particularly pronounced with the material for the capital. Unfortunately, it is not possible to do anything about these issues and they must simply be borne in mind in what follows.

A final important point is that inscriptions can be regarded as a reflection of two different, but related, kinds of information. The total number of inscriptions that were made can be regarded as an index of how many events were commemorated (i.e. how many different topics there are), while the total numbers of words in each inscription can be regarded as an index of how much detail was given about each event (i.e. how much content there is for each topic). This means that, in an ideal situation, one would want to be able to measure both the total number of inscriptions associated with each site and the total lengths of these inscriptions. One of the most important obstacles to doing this, however, is the fact that so many of these inscriptions are fragmentary or not complete, meaning that it is obviously not possible to estimate their lengths. Another complication is caused by the large numbers of abbreviations that are used in inscriptions, which means that their length is not necessarily a true reflection of the richness of their content. This is especially true of biographical material.

For the purposes of this article, I have therefore only counted the numbers of inscriptions that are associated with each site and have not attempted to estimate the total lengths of each inscription. There are strong reasons, however, for suspecting that the distribution of the lengths of these inscriptions would be log-normal, on the basis of comparison with modern sources that are more amenable to measurement, such as the material on internet sites. As recent research has shown, for example, the total lengths of the posts left by users on internet sites often conform to a log-normal distribution, regardless of their language, subject, or platform (Sobkowicz and others 2013). In addition, other studies suggest that there is a strong relationship between the numbers of unique words in a piece of text and the length of that text, so the length of a text equates to the richness of its content (this is known as Heaps' law) (Altmann and Gerlach 2016). This means that one would expect the total numbers and the total lengths of inscriptions to track each other relatively well, with the former being a decent proxy for the latter.

To link inscriptions and sites, I began by using a recent catalogue of ancient cities and towns to compile a list of relevant sites (Hanson 2016). I then looked up each site in the *EDCS*, using a combination of the ancient and modern names of sites to match them to the find-spots of inscriptions, bringing in their ancient provinces or modern countries as supporting evidence if necessary. I then made a count of the number of inscriptions recorded for each site. Since this database is mainly concerned with Latin, rather than Greek, inscriptions, I have concentrated on the regions that predominantly spoke this language, focusing on Italy and the western provinces, rather than the eastern provinces (although I have not deliberately removed any of the inscriptions with non-Latin and Greek characters).[4] I have not filtered inscriptions by date, however, since it is not possible to identify which inscriptions were on display at any one time. Although I have collected data about the capital, I have not included it in the following analysis, given that it is an extreme outlier from the relationship (in the technical, rather than the colloquial sense).[5] To locate inscriptions, I have used the coordinates given in the catalogue (Hanson 2016). Meanwhile, to date them, I have used the same date ranges as the *EDCS*. As noted above, however, one of the issues with these date ranges is that they are often quite broad, which means that using a simple start, middle, or end point to date them might not give a sufficiently nuanced impression of how they changed over time. I have therefore used the same method that Wilson (2009) has used on shipwrecks to estimate the probability that each inscription dates to each year, which can then be summed to give a new series.[6] To do this, I began by using the start and end dates of each inscription to calculate their date ranges and then used them to estimate the likelihood that each inscription existed in each year (this means an inscription dated to AD 1 to 2 is represented by 0.5 in AD 1 and 0.5 in AD 2, and so on). I have used a uniform distribution to model this, since our uncertainty about the dates that these inscriptions were created is essentially epistemic, rather than aleatory. The results are shown in Figure 7.1.

I then used the same method for estimating the populations of sites as elsewhere, which is based on their inhabited areas and densities (this avoids the problem of circularity that has appeared in earlier work, given that the numbers of inscriptions have sometimes been used to estimate the populations of sites) (Wilson 2011; Hanson 2016). Although these

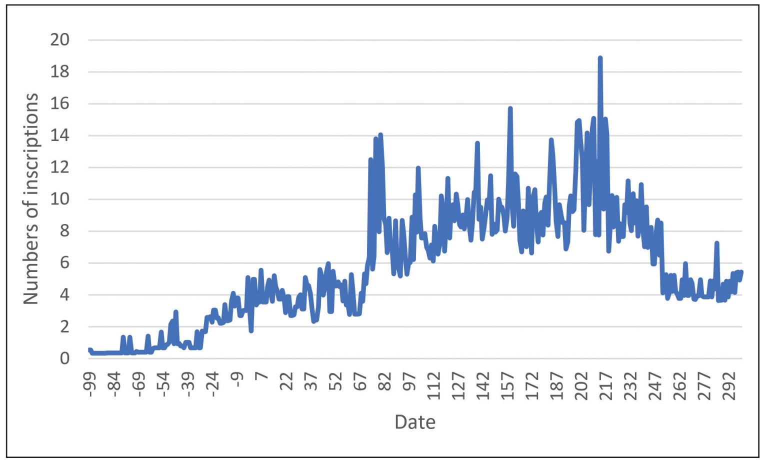

Figure 7.1. The numbers of inscriptions over time (summed probabilities per annum). These are generated using the same method as Wilson (2009) has used on shipwrecks. All of the inscriptions that can be dated have been included, rather than just those associated with sites, to make the results comparable with other proxies.

figures are approximate, they do seem to be sufficiently reliable to allow us to explore the relationships between the sizes of sites and their various attributes, as witnessed by recent studies of the densities of sites, the sizes of mixing spaces, such as *fora* and *agorai*, the dimensions of street networks, the widths of city gates, and the seating capacities of theatres and amphitheatres (Hanson and Ortman 2017; 2020; Hanson, Ortman, and Lobo 2017; Hanson and others 2019; Hanson 2020). It is also important to note that it is only necessary to estimate the order of magnitude of each site, given that the effects that we are interested in below are multiplicative. I have also assumed that most of the inscriptions that are associated with sites coincide with the maximum extension of each settlement. This is supported by the result of the probabilistic method for estimating the numbers of inscriptions over time referred to above, which shows that most of them were set up in the first or second century AD (Fig. 7.1). I have therefore used the cumulative numbers of inscriptions per site in what follows, rather than the estimated number in any given year or range of years, although using more refined date ranges would be an interesting matter for future research. I then assessed the relationship between the number of inscriptions associated with each site and the estimated population of each site using a common form of regression analysis known as ordinary least squares regression.[7]

4 These provinces include Africa Proconsularis, Alpes Cottiae, Alpes Graiae et Poeninae, Alpes Maritimae, Baetica, Britannia, Corsica et Sardinia, Dalmatia, Gallia Aquitania, Gallia Belgica, Gallia Lugdunensis, Gallia Narbonensis, Germania Inferior, Germania Superior, Hispania Tarraconensis, Italia (including I Latium and Campania, II Apulia et Calabria, III Lucania et Brutii, IV Samnium, V Picenum, VI Umbria and Ager Gallicus, VII Etruria, VIII Aemilia, IX Liguria, X Venetia et Histria, and XI Transpadana), Lusitania, Mauretania Caesariensis, Mauretania Tingitana, Noricum, Numidia, Pannonia Inferior, Pannonia Superior, Raetia, and Sicilia.

5 An outlier is generally defined as any observation that is more than two standard deviations away from the mean. 119,350 inscriptions are associated with Rome, while the standard deviation of the entire set of data is 5,124 and the mean is 571, therefore, any observation over 10,819 is an outlier. The capital is the only site that qualifies as such.

6 For more details about the background of this method and suggestions about how it could refine future research, see Johnson 2004; Broux 2019; and Crema and Kobayashi 2020.

7 This is feasible because $y = bx^m$ and $\log y = m \log x + \log b$ are equivalent expressions (Bettencourt 2013; Ortman and others 2014; Hanson and Ortman 2017; 2020; Hanson and others 2019). This means that the exponent and pre-factor of the

Table 7.1. The average numbers of inscriptions per site per bin.

Bin	Mean	Median
1.5 to 2	11	11
2 to 2.5	54	39
2.5 to 3	104	53
3 to 3.5	178	88
3.5 to 4	377	172
4 to 4.5	672	313
4.5 to 5	1327	721
5 to 5.5	4656	4656
5.5 to 6	119,350	119,350

Table 7.2. The total numbers of sites with at least one inscription, and the percentage of sites with at least one inscription per bin.

Bin	Number of sites	Number of sites with at least one inscription	Percentage (%)
1.5 to 2	2	2	100
2 to 2.5	14	11	79
2.5 to 3	82	76	93
3 to 3.5	190	174	92
3.5 to 4	194	189	97
4 to 4.5	89	83	93
4.5 to 5	19	16	84
5 to 5.5	2	2	100
5.5 to 6	1	1	100
Total	593	554	93

Finally, since there is a large amount of variation in the number of inscriptions associated with each site, I have used a standard econometric technique and grouped sites into a series of equally sized bins (Stock and Watson 2019). As Geoffrey West (2017, 387–88) has noted, this is a common strategy when confronted with big datasets with large fluctuations, since this approach allows us to overcome the variation in the data, revealing the underlying pattern. A good example of how this approach has been used is recent studies of the relationships between the sizes of companies and their net income, gross profit, total assets, and sales. As West (2017, 389) notes, although the results are coarse-grained, they nonetheless do a good job of describing the data and follow a good straight line. Interestingly, one of the reasons for the large variance in these data seems to be that they cover an unusually long time period, at least for contemporary data, from the 1950s to the 2000s. This is reminiscent of the data discussed in this article, which are also derived from a long time span. This method is useful, therefore, for looking at the relationship between the sizes of sites and their various attributes, since it essentially allows us to control for variation between sites, revealing the otherwise hidden effects of scale. In the context of inscriptions, this might be especially important, since one would expect there to be a general relationship between them, which is masked by the large amount of variation between sites resulting from the differing epigraphic habits discussed above and variation in the quality and quantity of information available for each site. As noted above, although one would expect there to be a great deal of variation from site to site, as a result of both different amounts of evidence and different epigraphic traditions, it is possible to use this method to average out and overcome these biases, so long as a large sample is used. I have therefore used the same method as West and binned the data by the estimated population of sites, before calculating the average of the number of inscriptions that are associated with each site per bin (Table 7.1). These bins are based on log-transformed data, which means that the first bin ranges from 1 to 10, the second from 10 to 100, the third from 100 to 1000, and so on (West 2017, 389). I have used eight bins, each of which spans half a logarithm, with the result that there are two bins for each order of magnitude. As can be seen in Table 7.2, this approach provides us with data for a minimum of just under 80 per cent of the sites in each bin. The percentage for most of these bins, however, is actually much higher — between 90 and 100 per cent. This means that it is possible to generate a relatively consistent average figure of the number of inscriptions associated with sites across bins.

One of the drawbacks of this approach is that it does not allow us to calculate the extent to which individual sites deviate from the overall relationship, a measure that is known as a residual. These residuals can be simulated, however, by using the equation for the best-fit line of the relationship between the average number of inscriptions per site per bin and their estimated populations to generate an expectation for each site. It is then possible to subtract the expected figure from the observed figure to produce a value that is equivalent to a residual. I have used the median, rather than the mean, given that

power function can be estimated by calculating the slope and y-intercept of the best-fit linear function, so long as it is based on log-transformed values. I have not controlled for heterogeneity.

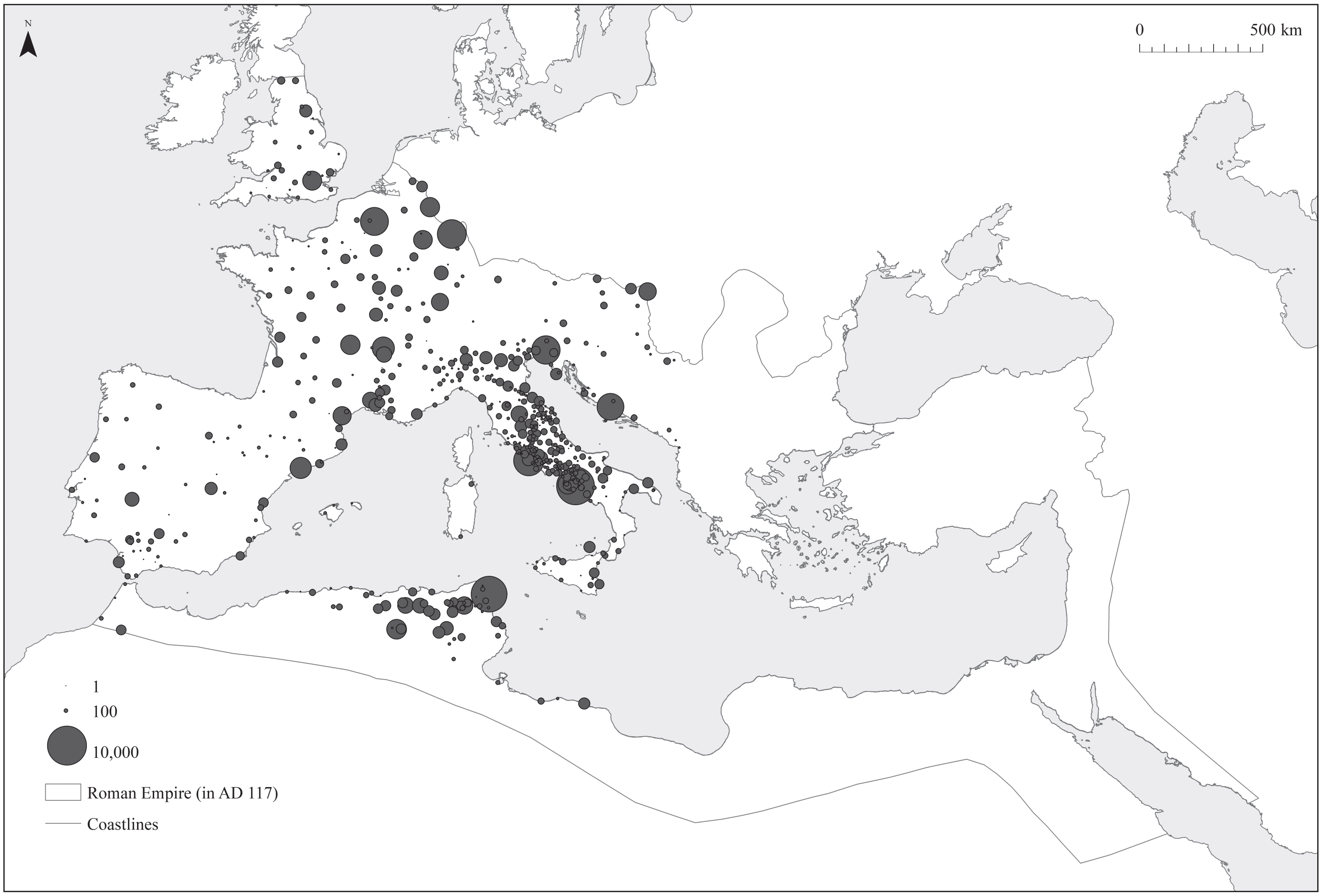

Figure 7.2. The numbers of inscriptions associated with each site. Symbols are proportional to the numbers of inscriptions. The capital is not shown. Map by author.

this provides a more balanced number of positive and negative residuals. Here, I also suggest a slight improvement on earlier work by investigating the distribution of the total numbers of inscriptions and these residuals. This uses a common approach, which is to investigate the extent to which sites with similar attributes are found near to one another (this is known as autocorrelation) (Wheatley and Gillings 2002; Conolly and Lake 2006). This can be done by using a standard tool in ArcGIS, which produces a single index where different values indicate that sites are random, dispersed, or clustered (Table 7.4).

Results and Discussion

In total, 316,507 inscriptions have been used in the analysis, which are associated with 554 sites (Fig. 7.2). These sites are drawn from throughout the settlement hierarchy and cover several orders of magnitude, from almost a million to less than a thousand inhabitants, and come from throughout Italy and the western provinces. The number of inscriptions is interesting in its own right, since it demonstrates that the majority, although not an overwhelming majority, of the inscriptions that are known from the ancient world are associated with urban environments (62.8 per cent based on the current tallies of inscriptions in the *EDCS* at the time the research was undertaken). In addition, these figures illustrate both the importance of inscriptions to urban life and the scale of the material that survives, given that there are only 39 sites in Italy and the western provinces for which there is no evidence, out of the total of 593 that are known to have existed (Hanson 2016).

As can be seen in Figure 7.3, although there is a wide range of variation in the number of inscriptions per site, the overall distribution is dominated by the capital. These numbers cover a wide range, from 119,350 to 1, the average being 571 (a complete list of the numbers of inscriptions associated with each site is given in the supplementary material).[8] The distribution of inscriptions also conforms well

8 Supplementary material: <https://doi.org/10.1484/A.14821884>.

Table 7.3. The results of regressing the average numbers of inscriptions per site per bin against the estimated populations of sites. All regressions are done using ordinary least squares regression on log-transformed values. The capital, Rome, has been excluded from the analysis, as it is an extreme outlier.

Dependent variable	Number of cases	Exponent (95% CI)	Pre-factor (95% CI)	R^2	Significance (P-value)
Mean	8 bins (554 sites)	0.672 (0.588–0.756)	1.122 (0.552–2.279)	0.976	<0.0001
Median	8 bins (554 sites)	0.654 (0.534–0.774)	0.800 (0.288–2.218)	0.950	<0.0001

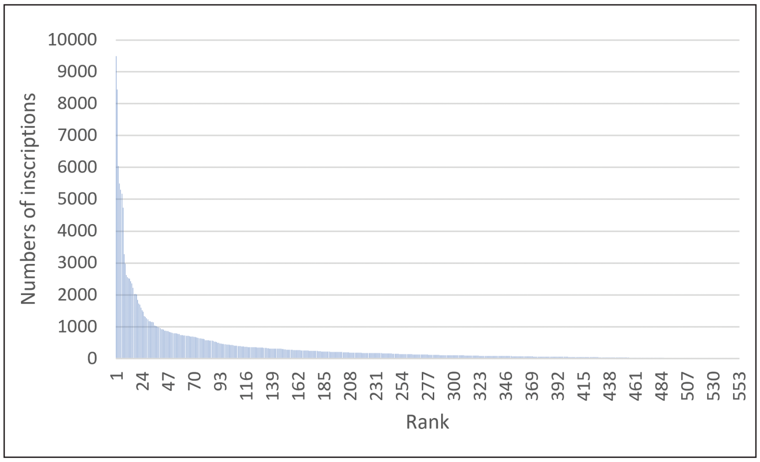

Figure 7.3. The numbers of inscriptions associated with selected sites. The capital is not shown to make the variation between the remaining sites visible.

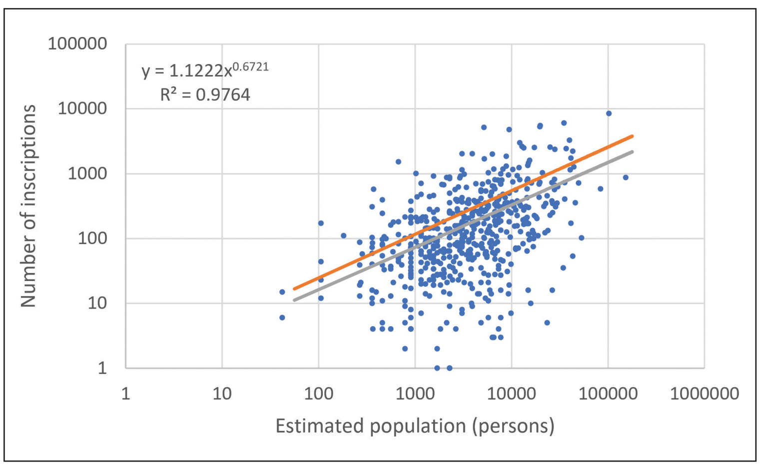

Figure 7.4. The relationship between the number of inscriptions per site and the estimated population. The blue dots represent the number of inscriptions per site, while the orange line is the best-fit line for the mean number of inscriptions per site per bin, and the grey line is the best fit line for the median number of inscriptions (again per site per bin). The inset equation and R^2 refers to the binned data. Both scales are logarithmic. The capital, Rome, is not shown, as it is an extreme outlier.

to our expectations. In fact, although most of them are concentrated in Italy, there are also large numbers in North Africa and Gaul and Germany, following the distribution of urban centres.

The results show that there is a consistent relationship between the average numbers of inscriptions per site per bin and their estimated populations, which is entirely in keeping with both wider theoretical and empirical expectations and the new model discussed above (Table 7.3 and Fig. 7.4). This furnishes us with the first concrete evidence, at least for the ancient world, that the numbers of artefacts vary with the sizes of settlements. This relationship is not linear, however, and the total numbers of inscriptions associated with sites grow more slowly than their estimated populations. In other words, there was a decreasing number of inscriptions per capita as the sizes of cities increased. This is reflected in the slope, or exponent, of the relationship, which is about two-thirds (0.67). This suggests that it might be best to regard inscriptions as a form of infrastructure for conveying information and that the numbers of inscriptions should not, by themselves, be regarded as a reflection of increasing wealth or disposable income. As noted above, this result can be explained by the fact that one would expect larger cities to have more social interactions and higher population densities, so one would expect each inscription to have been viewed and discussed more frequently, on average, as the number of residents grew. This means that fewer inscriptions per capita would have been needed in larger cities to fulfil the same role. These results are striking, therefore, since they show that a form of evidence that was undoubtedly intended to convey information increases with the estimated populations of sites at the same rate as other forms of infrastructure in other contexts. This not only supports the idea that cities can be characterized as complex systems, which depend on flows of material, energy, and information, but also that the roles of settlements as generators of and vessels for information are not unique to the modern world.

It is important to remember, however, that these results are based on the total numbers, rather than the total lengths of inscriptions. These results might therefore suggest a slightly more nuanced hypothesis for how the interplay of information and wealth affects the numbers and lengths of inscriptions per site, where the raw numbers of inscriptions might be seen as a reflection of the numbers of events deemed important enough to be recorded, and their lengths might be seen as a reflection of the amount of detail that it was decided to give about each one. This means that the extent of any disparity between the expected and observed lengths of inscriptions in a given context could be a more sensitive reflection of disposable income.

These results support earlier findings, which suggested a similar relationship between the numbers of inscriptions associated with sites and their estimated population, based on a sample of inscriptions associated with sites with known professional associations (Hanson, Ortman, and Lobo 2017, table 2). This research suggested an exponent of about 0.64 and a pre-factor of -0.34, with an R^2 of 0.58, although the size of the sample was much smaller and only included 210 sites. These results also echo the results for other forms of infrastructure, given that inscriptions increase at the same rate as not only the sizes of the inhabited areas of cities themselves, but also the dimensions of mixing spaces, such as *fora* and *agorai*, and street networks (Hanson and others 2019). It is interesting to note, however, that inscriptions increase faster than the seating capacities of theatres and amphitheatres (Hanson and Ortman 2020). As noted above, this makes sense, given that one would expect the former to be a product of the social networks and densities of cities, while one would expect the latter to be a product of the transmission of information through settlements from person to person, resulting in different scale dependencies. Another possibility is that the information that was contained in inscriptions was less significant or relevant to most inhabitants than the news about what happened in theatres and amphitheatres, and was therefore slightly less likely to be repeated from person to person, although it was nonetheless important to urban life and played a vital role in social cohesion.

As noted above, although one would expect the main purpose of inscriptions to have been to convey socially important information, it is important to be extremely careful about using them as a proxy for the total amount of information that was contained within a settlement, given that they were not the only sources of information. Having said this, when taken in conjunction with the model described above, these results do raise the interesting possibility that the total amount of information that was contained in settlements increased faster than the estimated populations of sites. If this is correct, it would mean that, although larger sites did generate more information than their smaller counterparts, they were also associated with an increasing amount of information per capita. This would make sense, given that information is not rivalrous and can be shared by more than one individual at once. This might have significant implications for our view of urbanism, since recent work has underscored the crucial role that information has in allowing a city or a system of cities to evolve (Bettencourt 2014). More investigation is needed, however, before these kinds of conclusions can be drawn.

These results also show, however, that there is a great deal of variation in the number of inscriptions per site, which means that the amount of evidence that is available for each of them fluctuates significantly. This can be explained by a combination of variations in the tendency of individuals to record information in inscriptions between different periods and places and discrepancies in levels of preservation, recovery, and publication across sites — not to mention mismatches in the dates of inscriptions and estimated populations, and variation in the kind of material included in the count for each site. This might explain why it has been so difficult to use inscriptions to reconstruct the demographic conditions of the ancient world, as there is simply too much variation in the numbers of inscriptions associated with sites to use them as evidence for their numbers of inhabitants. In other words, although it is possible to identify an overall relationship between the numbers of inscriptions and the sizes of sites, it is not possible to use one to predict the other at the level of individual settlements.

These results suggest that both perspectives discussed in the introduction to this article are partly right, in the sense that, although there is a broad relationship between the numbers of inscriptions associated with sites and their estimated populations, which supports the notion of an overall 'epigraphic habit', there is a large amount of variance in this relationship, revealing significant differences in the epigraphic cultures of each site. These results might also help us to explain one of the main uncertainties about what inscriptions represent, since they allow us to see the very existence of a relationship as a reflection of the shared impulse of individuals to use inscriptions and the extreme amount of variation between sites as a reflection of differences in the tendencies of individuals to make some statement through the medium of an inscription across space and time. In other words, it is only by paying attention both to how much the numbers of inscriptions vary with the sizes of sites and how much individual sites deviate from the overall relationship that the central paradox at heart of the epigraphic culture of the ancient world might be addressed: namely, the apparent ubiquity of inscriptions, alongside extreme variation.

There does not seem to be any obvious pattern in the distribution of residuals (Figs 7.5 and 7.6). This supports the idea that there is a lot of random noise in the data, which has been controlled for

Table 7.4. The results of investigating the spatial autocorrelation of the numbers of inscriptions associated with sites and the residuals of the inscription-estimated population relationship. Moran's index is a measure of the extent to which the features are random, dispersed, or clustered, while the z-score and p-value are measures of the significance of the results. The results suggest that, although the distribution of the numbers of inscriptions associated with sites is indistinguishable from random, the distribution of the residuals of the inscription-estimated population relationship is clustered.

Variable	Interpretation	Moran's index	Z-score	P-value
Numbers of inscriptions	Random	-0.006	-1.076	0.282
Residuals	Clustered	0.046	4.571	<0.0001

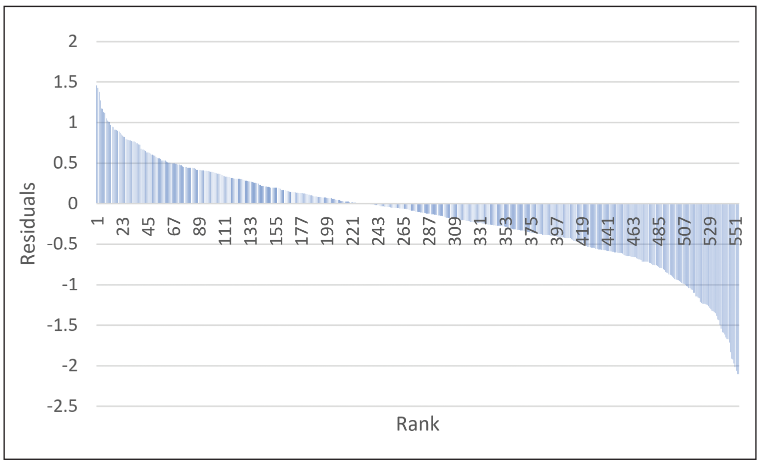

Figure 7.5. The residuals of the inscription-estimated population relationship. These are derived by using the best-fit line for the relationship between the average numbers of inscriptions per site per bin and their estimated populations to create an expectation for each site and then subtracting the expected value from the observed value. Here, the median is used, rather than the mean. These residuals are arranged from largest to smallest, from positive to negative.

by binning inscriptions. The results of testing for autocorrelation, however, also show that, although there does not seem to be any relationship between the distribution of sites and the basic numbers of inscriptions that are associated with them (which are indistinguishable from random), there is a relationship between the distribution of sites and the sizes of their residuals (which are clustered) (the results of using the standard tool in ArcGIS are shown in Table 7.4). This suggests that an interesting avenue for future work might be to investigate the distribution of the residuals of this and similar relationships in more detail. Although one would expect the scale of sites to have the greatest impact on their characteristics, one might also expect their relationship with their immediate neighbours to be one of the most important factors behind the extent to which they under- or overperformed relative to their size. Investigating this in more detail could be an important subject for future research.

In addition, there is a crude relationship between the sizes of the residuals and the civic statuses of sites. The average residual for provincial capitals, which is 0.43, is larger than the average figures for both *coloniae* and *municipia*, which are both about 0.06. This suggests that the sites with the highest civic statuses were more likely to have had a larger number of inscriptions than might be anticipated based on their size. This supports the results from earlier studies, which focused on the seating capacities of theatres and amphitheatres, revealing a similar, although not identical, pattern (Hanson and Ortman 2020).

Note also that the sites from different regions, such as Britannia, Gaul and Germany, and Italy, are evenly scattered throughout the relationship, which suggests that the amount of *instrumentum domesticum* that has been included has little impact on the results (this is because the inscriptions from different regions contain different amounts of this material). This could be investigated in more detail in future work, such as quantifying the different amounts of different kinds of material on a site-by-site basis.

It is important to remember, however, that the approach that has been advocated here is only intended to capture the average conditions across settlements and that individual sites will conform to these relationships to various degrees. As a result, rather than discounting the possibility of variation between sites, this article hopes to have opened up several avenues for future research by suggesting that, although there is an overall relationship between the average numbers of inscriptions per site and the estimated populations of sites, there is also significant variation from site to site that requires further explanation. As noted above, this variation could be the result of a number of factors including not only differences in the amount of each site that has been excavated or the numbers of inscriptions that have been recorded and published, but also differences in the epigraphic traditions of sites, including differences in the extent to which different languages were used as the dominant form of communication, etc. An important next step for scholars might therefore be to attempt to quantify these secondary factors on a site-by-site basis and examine whether there is any correlation between the sizes of these variables and the sizes of the residuals mentioned above, expanding on the discussion of the links between the residuals and civic statuses of sites above. The strength of these correlations could then be used to assess the relative importance of these factors, fur-

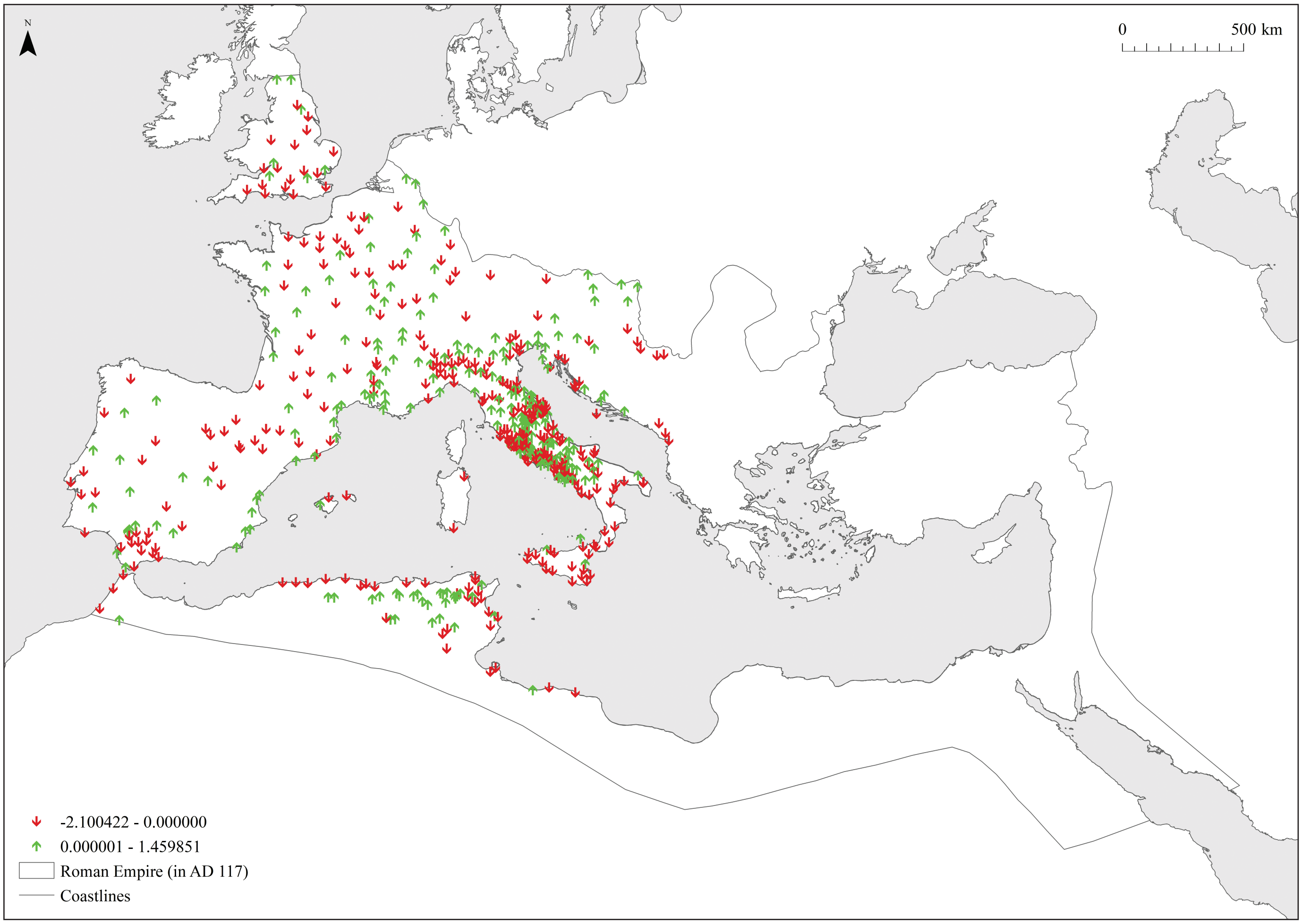

Figure 7.6. The residuals of the inscription-estimated population relationship. For more details about how these residuals are calculated, refer to the caption of Figure 7.5. Map by author.

ther helping to clarify our understanding of the epigraphic conditions of the ancient world.

Apart from this, another important question is how much the results would change if it had been possible to consider a larger amount of material from a wider selection of languages. As noted above, although one would not necessarily expect this to affect the slope of the relationship, this might account for some of the variation in the extent to which individual sites conform to or deviate from it (i.e. for the variation in the residuals), potentially resulting in a less noisy relationship. This could be tested by integrating other resources. In addition, it would be interesting to investigate the effects of including the lengths of inscriptions to assess whether they follow the same pattern, as well as to investigate whether different kinds of inscriptions, such as funerary or building inscriptions, scale at different rates. One hypothesis is that the numbers of epitaphs would increase faster than the numbers of building inscriptions, given their different audiences and 'end users'.

As Wilson (2011, 163–67) has argued in recent work, it is possible to use the numbers of inscriptions associated with the construction, maintenance, and repair of buildings as a crude reflection of the wider construction industry and therefore as a reflection of economic development. This raises the question of whether it is possible to use the same, or a similar, subset to investigate how the amount of documented building work changed with the sizes of cities, with important implications for our view of ancient economic life. In addition, it would be possible to investigate how much the contents of inscriptions change with the sizes of sites and, in particular, whether inscriptions in larger sites express different concerns. This could be achieved, for example, by not only searching inscriptions for certain words and phrases, but also by using topic models to detect groups of inscriptions that contain similar words or phrases, identify categories of inscriptions, and assess the extent to which these categories intersect. A preliminary analysis shows that the total numbers of inscriptions that contain the phrase *D(is) M(anibus)* (to the gods of the underworld) seem to grow more slowly than all kinds of inscriptions, suggesting that there is more to be uncovered. It would also be profitable to re-evaluate other kinds of artefacts, such as coinage and ceramics, which have traditionally been assumed to have had little or no relationship with the sizes of sites, assuming that sufficient data are available to do this. Finally, this approach might also be useful in other contexts with large volumes of inscriptions, even if they are pictographic, such as Central America.

Conclusions

Although scholars have traditionally been extremely pessimistic about the value of quantifying ancient inscriptions and about the idea that inscriptions reflect meaningful differences between sites, in this article I have shown that there is a consistent relationship between the numbers of inscriptions associated with sites and their estimated populations, which entirely conforms to wider theoretical and empirical expectations, suggesting that the roles of settlements as generators of and vessels for information are not unique to the modern world. These results suggest that, although larger cities did generate a larger number of inscriptions than their smaller counterparts, inscriptions increased more slowly than their estimated populations. This indicates that inscriptions acted as a kind of information infrastructure and were not, by themselves, a simple reflection of increasing wealth or disposable income. I have suggested that both of the perspectives discussed in the introduction are partly right, in the sense that, although there is a broad relationship between the numbers of inscriptions associated with sites and their estimated populations, which supports the notion of an 'epigraphic habit', there is a large amount of variation in this relationship, revealing significant differences in the epigraphic cultures between sites. This result has significant implications, given the importance of epigraphic material to our understanding of various aspects of the ancient world. It might also have an important bearing on our view of how information spread throughout settlements, suggesting fruitful avenues for future research.

Abbreviations

CIL *Corpus inscriptionum latinarum*.

EDCS *Epigraphik-Datenbank Clauss / Slaby*.

EDH *Epigraphische Datenbank Heidelberg*.

Works Cited

Altmann, E. G. and M. Gerlach. 2016. 'Statistical Laws in Linguistics', in M. Degli Esposti, E. Altmann, and F. Pachet (eds), *Creativity and Universality in Language* (Cham: Springer), pp. 7–26.

Beltrán Lloris, F. 2014. 'The "Epigraphic Habit" in the Roman World', in C. Bruun and J. Edmondson (eds), *The Oxford Handbook of Roman Epigraphy* (Oxford: Oxford University Press), pp. 131–48.

Bettencourt, L. M. A. 2013. 'The Origins of Scaling in Cities', *Science*, 340: 1438–41.

——. 2014. 'Impact of Changing Technology on the Evolution of Complex Informational Networks', *Proceedings of the IEEE*, 102: 1878–91.

Bettencourt, L. M. A. and G. B. West. 2010. 'A Unified Theory of Urban Living', *Nature*, 467: 912–13.

Bodel, J. 2001. *Epigraphic Evidence: Ancient History from Inscriptions* (London: Routledge).

Broux, Y. 2019. 'An Improved Weighed Dates Method for Ancient People', *Ancient Society*, 49: 103–21.

Bruun, C. and J. Edmondson (eds). 2015. *The Oxford Handbook of Roman Epigraphy* (Oxford: Oxford University Press).

Cherry, D. 1995. 'Re-figuring the Roman Epigraphic Habit', *Ancient History Bulletin*, 9: 143–56.

Conolly, J. and M. Lake. 2006. *Geographical Information Systems in Archaeology* (Cambridge: Cambridge University Press).

Cooley, A. (ed.). 2012. *The Cambridge Manual of Latin Epigraphy* (Cambridge: Cambridge University Press).

Crema, E. R. and K. Kobayashi. 2020. 'A Multi-Proxy Inference of Jōmon Population Dynamics Using Bayesian Phase Models, Residential Data, and Summed Probability Distribution of ^{14}C Dates', *Journal of Archaeological Science*, 117: 105136.

Duncan-Jones, R. P. 1982. *The Economy of the Roman Empire: Quantitative Studies*, 2nd edn (Cambridge: Cambridge University Press).

Hanson, J. W. 2016. *An Urban Geography of the Roman World, 100 BC to AD 300* (Oxford: Archaeopress).

——. 2020. 'Using City Gates as a Means of Estimating Ancient Traffic Flows', *PLoS ONE*, 15.2: e0229580.

Hanson, J. W. and S. G. Ortman. 2017. 'A Systematic Method for Estimating the Populations of Greek and Roman Settlements', *Journal of Roman Archaeology*, 30: 301–24.

——. 2020. 'Reassessing the Capacities of Entertainment Structures in the Roman Empire', *American Journal of Archaeology*, 124: 417–40.

Hanson, J. W., S. G. Ortman, and J. Lobo. 2017. 'Urbanism and the Division of Labour in the Roman Empire', *Journal of the Royal Society Interface*, 14 <https://doi.org/10.1098/rsif.2017.0367>.

Hanson, J. W., S. G. Ortman, L. M. A. Bettencourt, and L. C. Mazur. 2019. 'Urban Form, Infrastructure, and Spatial Organisation in the Roman Empire', *Antiquity*, 93.369: 702–18.

Harvey, B. K. 2004. *Roman Lives: Ancient Roman Life as Illustrated by Latin Inscriptions* (Newburyport: Focus).

Johnson, I. 2004. 'Aoristic Analysis: Seeds of a New Approach to Mapping Archaeological Distributions through Time', in K. F. Ausserer, W. Börner, M. Goriany, and L. Karlhuber-Vöckl (eds), *[Enter the Past]: The E-Way into the Four Dimensions of Cultural Heritage; CAA 2003: Computer Applications and Quantitative Methods in Archaeology; Proceedings of the 31st Conference, Vienna, Austria, April 2003* (Oxford: Archaeopress), pp. 448–52.

Keppie, L. 1991. *Understanding Roman Inscriptions* (London: Batsford).

Lobo, J., L. M. A. Bettencourt, M. E. Smith, and S. G. Ortman. 2020. 'Settlement Scaling Theory: Bridging the Study of Ancient and Contemporary Urban Systems', *Urban Studies*, 57.4: 731–47.

MacMullen, R. 1982. 'The Epigraphic Habit in the Roman Empire', *The American Journal of Philology*, 103: 233–46.

Mann, J. C. 1985. 'Epigraphic Consciousness', *Journal of Roman Studies*, 75: 204–06.

Meier, R. L. 1962. *A Communications Theory of Urban Growth* (Cambridge, MA: MIT Press).

Meyer, E. A. 1990. 'Explaining the Epigraphic Habit in the Roman Empire: The Evidence of Epitaphs', *Journal of Roman Studies*, 80: 75–96.

Ortman, S. G., A. H. F. Cabaniss, J. O. Sturm, and L. M. A. Bettencourt. 2014. 'The Pre-history of Urban Scaling', *PLoS ONE*, 9.2: e87902.

Ortman, S. G., M. E. Smith, J. Lobo, and L. M. A. Bettencourt. 2020. 'Why Archaeology Is Necessary for a Theory of Urbanization', *Journal of Urban Archaeology*, 1: 151–67.

Saller, R. P. and B. D. Shaw. 1984. 'Tombstones and Roman Family Relations in the Principate: Civilian, Soldiers, and Slaves', *Journal of Roman Studies*, 74: 124–56.

Schläpfer, M., L. M. A. Bettencourt, S. Grauwin, M. Raschke, R. Claxton, Z. Smoreda, G. B. West, and C. Ratti. 2014. 'The Scaling of Human Interactions with City Size', *Journal of the Royal Society Interface*, 11 <https://doi.org/10.1098/rsif.2013.0789>.

Smith, M. E. 2018. 'Energized Crowding and the Generative Role of Settlement Aggregation and Urbanisation', in A. Gyucha (ed.), *Coming Together: Comparative Approaches to Population Aggregation and Early Urbanization* (Albany: State University of New York Press), pp. 37–58.

Sobkowicz, P., M. Thelwall, K. Buckley, G. Paltoglou, and A. Sobkowicz. 2013. 'Lognormal Distributions of User Post Lengths in Internet Discussions: A Consequence of the Weber-Fechner Law?', *EPJ Data Science*, 2: article 2 (2013).

Stock, J. H. and M. W. Watson. 2019. *Introduction to Econometrics*, 4th edn (Harlow: Pearson).

West, G. B. 2017. *Scale: The Universal Laws of Life and Death in Organisms, Cities, and Companies* (London: Weidenfeld & Nicolson).

Wheatley, D. and M. Gillings. 2002. *Spatial Technology and Archaeology: The Archaeological Applications of GIS* (London: Taylor & Francis).

Wilson, A. I. 2009. 'Approaches to Quantifying Roman Trade', in A. K. Bowman and A. I. Wilson (eds), *Quantifying the Roman Economy: Methods and Problems* (Oxford: Oxford University Press), pp. 213–49.

——. 2011. 'City Sizes and Urbanisation in the Roman Empire', in A. K. Bowman and A. I. Wilson (eds), *Settlement, Urbanisation, and Population* (Oxford: Oxford University Press), pp. 161–95.

——. 2014. 'Quantifying Roman Economic Performance by Means of Proxies: Pitfalls and Potential', in F. de Callataÿ (ed.), *Quantifying the Greco-Roman Economy and Beyond* (Bari: Edipuglia), pp. 147–67.

Woolf, G. 1996. 'Monumental Writing and the Expansion of the Roman Society in the Early Empire', *Journal of Roman Studies*, 86: 22–39.