About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world’s best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org
WHERE DO CITIES COME FROM AND WHERE ARE THEY GOING TO? MODELLING PAST AND PRESENT AGGLOMERATIONS TO UNDERSTAND URBAN WAYS OF LIFE

Topic Editors:
Francesca Fulminante, University of Bristol, United Kingdom
John William Hanson, University of Colorado Boulder, United States
Scott G. Ortman, University of Colorado Boulder, United States
Luis M. A. Bettencourt, University of Chicago, United States

Over the last decade, there has been a surge of interest in urbanization and economic development, sparked by the realization that making urban life sustainable is one of the greatest challenges facing us in the 21st century (this is now one of the core sustainable development goals of the United Nations). This has exerted considerable pressure on researchers to come up with more scientific ways of studying urbanism and economic activity over the long run, which has resulted not only in the development of new theoretical frameworks, but also in the collection of vast amounts of data from a range of settings.

This has led to the realization that, although there are significant differences between settlements in different settings, there are nonetheless important regularities and commonalities between a diverse group of settlements in range of geographical and historical contexts, including both ancient and modern ones. This suggests that a common feature of settlements is their ability to generate increased social connectivity, greater division of labour and specialization, and enhanced technological invention and innovation, albeit with costs to levels of equality, quality of life, and standards of living, as well as impacts on the environment, which cannot be separated from the emergence of confederations and states and the creation of settlement systems, hierarchies and networks.

We believe that this field of enquiry now stands at a critical juncture. Although it is now feasible to talk about many aspects of ancient and modern urbanism with relative confidence, such as the numbers of cities or their sizes, much of the discussion of these themes within historical and archaeological circles has been on a discursive or qualitative level, while it is often difficult to harmonize the different models that have been applied to date into a consistent empirical and theoretical framework. A new approach to settlements throughout different contexts should now be within our grasp, however, thanks to both the ease with which information can be disseminated and the facilities that recent developments in IT offer us to model, analyse, and statistically test data.

# Table of Contents

04 **Editorial: Where Do Cities Come From and Where Are They Going To?**  
*Modelling Past and Present Agglomerations to Understand Urban Ways of Life*  
Francesca Fulminante

10 **Socio-Material Archaeological Networks at Çatalhöyük a Community Detection Approach**  
Camilla Mazzucato

35 **The Origins of Trypillia Megasites**  
John Chapman, Bisserka Gaydarska and Marco Nebbia

55 **Cities Through the Ages: One Thing or Many?**  
Michael E. Smith and José Lobo

65 **More Real Than Ideal: Household and Community Diversity at Metapontum, South Italy**  
Andrew H. F. Cabaniss

76 **Trajectories to Low-Density Settlements Past and Present: Paradox and Outcomes**  
Roland Fletcher

97 **Modeling the Rise of the City: Early Urban Networks in Southern Italy**  
Lieve Donnellan

116 **A New Kind of Relevance for Archaeology**  
Scott G. Ortman

129 **Quantitative Methods for the Comparative Analysis of Cities in History**  
Luís M. A. Bettencourt and José Lobo

137 **Ancient City, Universal Growth? Exploring Urban Expansion and Economic Development on Rome’s Eastern Periphery**  
Matthew J. Mandich

154 **Urbanity as a Process and the Role of Relative Network Properties—A Case Study From the Early Iron Age**  
Oliver Nakoinz, Michael Bilger and David Matzig

170 **Patterns of Etruscan Urbanism**  
Simon Stoddart, Alessio Palmisano, David Redhouse, Graeme Barker, Giorgia di Paola, Laura Motta, Tom Rasmussen, Troy Samuels and Robert Witcher
Editorial: Where Do Cities Come From and Where Are They Going To? Modelling Past and Present Agglomerations to Understand Urban Ways of Life

Francesca Fulminante 1, 2*

1 Department of Anthropology and Archaeology, Bristol University, Bristol, United Kingdom, 2 Division of Humanities, University Roma Tre, Rome, Italy

Keywords: urbanism, Italy, Europe, world archaeology, comparative quantitative perspective

Editorial on the Research Topic

Where Do Cities Come From and Where Are They Going To? Modelling Past and Present Agglomerations to Understand Urban Ways of Life

Urbanism in the past and present remains hotly debated in academia and the media. We can think of a series of successfully conducted projects in the last few years: for example, the Copenhagen Polis Centre project; the Reception of the City in Late Antiquity ERC project (Cambridge, UK); the ongoing UrbNet project (Aarhus, Denmark); the Social Reactors Project (Colorado USA). To these now the Dutch Universities OIKOS network can also be added, and if this was not enough the Guardian has recently launched a series "Guardian Cities" in the UK media. Yet fundamental questions such as "What is an ancient city? when can we say that a nucleated settlement has become a city? Why sometime a city prevails over others and why eventually it declines?"; are still widely open and lively debated question, that have not received a definitive answer yet especially with reference to central Italy, and Rome in particular.

The long-term trajectory of Rome is quite well-known and established from the early supremacy within Latium vetus in pre-historic and early historic times, to the emerging power in Italy, during the Republican period, and finally the dominance over the Empire, in the first few centuries of our Era before the final collapse around the end of the fourth century AD. However, the contributory factors and the determinants of this trajectory, which took “a slightly shabby Iron Age village” to become the “undisputed hegemon of the Mediterranean” are still very much questioned 1. In this editorial I will discuss features of urbanism/urbanization by presenting the current debate on the ancient city, also with reference to the recent Cambridge University Press book by Arjan Zuiderhoek 2, which summarizes and discusses extensively previous approaches. Then I will discuss the contribution of this special Research Topic and I will indicate further possible points of debate.

Already in the Bronze Age, but more commonly with the advent of the Iron Age, within the Near East, in Europe but also in the Americas, many regions become organized in small independent political units, generally defined as city-states 3. Since the classic work by Fustel de Coulanges, La Cité Antique, published in 1864 4, the debate on the characteristics and the origin of the ancient

---

1 See for example also the recent synthesis by Beard (2015), reviewed by the Wall Street Journal (quotations in brackets).
2 Zuiderhoek (2016). This book and the current debate is illustrated also in more detail in Fulminante (forthcoming).
4 Fustel de Coulanges (1864).
city has been immense, but the scholarly and at the same time agile book by Zuiderhoek, help us navigate into this dense and intricated subject. On one hand, Zuiderhoek discusses classical models of the ancient city, such as those elaborated by:

1) Fustel de Coulanges: based on a primordial, Indo-European notion of private property, originated in claims of land control and household possession through the cult of ancestors;
2) Max Weber: contrasting the modern-medieval city economy to the ancient household economy;
3) Moses Finley: conceptualizes the ancient city as a consumer city (greatly influenced by Max Weber) to explain the ancient world’s relative economic underdevelopment, in comparison with medieval and early modern Europe.

As observed by Zuiderhoek, in stressing the contrast with antiquity, all these three famous and influential models were interested in emphasizing the exceptionalism of Western European medieval cities, from which the unique development toward capitalism, the Industrial Revolution and modern liberal society would have emerged.

Besides these fundamental and influential models of the ancient city Zuiderhoek discussed all major models of urbanism developed by past and current scholarship, that can be summarized and integrated with further discussion as follow:

1. The demographic model can be based either on settlement size, with urban setting recognized above the threshold of 10,000 individuals or in the case of ancient cities, 5,000; the density/nucleation principle, according to which “cities are places where a certain energized crowding of people takes place”; or the demographic composition of the population with the alternative models of the “graveyard,” in which high urban mortality rates due to dirty and overcrowded environments, especially among infant/children, require immigration to explain urban growth) and “demographic transition” model, according to which higher fertility rates, led by early cessation of breastfeeding, could overweight high urban mortality rates, allowing for population survival and reproduction and eventually the demographic and economic growth.

2. More classic, the socio-economic model, characterizes urbanism by specialization of labor, social stratification and complementarity between the consumer city and the producing countryside, that is the market economy.

3. The model of urban environment and/or urban Landscape, based on the appearance of the ancient city, “with the presence of central squares or plazas, paved streets, defensive walls and gates, public architecture for religious, political or ceremonial/entertainment purposes and some element of town planning. It is perhaps in this sphere that the intuitive understanding of a settlement as ‘urban’ (we know it when we see it) is strongest."

4. The political model, according to which “Greek and Roman cities were political communities, which possessed the institutions required for autonomous collective decision-making.”

5. The ritual and identity model according to which cities were communities not only for full members of the political body (civitas) but a wider group of people, including women, children, freedmen, resident foreigners and slaves, that were effectively non or semi-citizens but would find unity and interactions in the comprehensive and inclusive action of the city rituals and festivals. While religion has often been connected to power as a mean of coercion and ideological control (Religio Instrumentum Regni), from ancient classical authors to Niccolò Macchiavelli’s treatise, Jörg Rüpke is developing a new dynamic way of looking at religion as a mean of actively creating power and the changes that led to early states societies.

To these models identified by Zuiderhoek, now has also to be added the “house society” model, originally developed by Claude Lévi-Strauss and since elaborated on by numerous scholars, also with reference to Mediterranean Bronze and Iron Age societies and to Central Italy, in particular. This model emphasizes the role of the family as an institution, with related anthropological and social practices such as marriages, hereditary rights etc. and seems to offer the missing link between egalitarian pre-urban societies and stratified and hierarchical urban developments, also being a key factor, in a dialectic manner, for the creation of state institutions. This view, reminiscent of Karl Marx and Friederich Engels perspectives, had already been

---

6See above note 4.
15 Elaborated by McLaren (1978) for Early Industrial London, the model has been used by Haydock et al. (2013) for Roman and Medieval Britain, but could have a much wider applicability, see Fulminante (2015).
16 For this model see Weber’s and Finley’s theories discussed above and further discussion in Zuiderhoek (2016), Chapter 3.
17 Zuiderhoek (2016), p. 6 and chapter 4. For a monumental approach to urbanism see Creekmore and Fisher (2014); with particular reference to Rome and Central Italy see Cifani (2008, 2010, 2014, 2018). Also contributions in Thomas and Meyers (2012). With reference to this approach new technological developments such as LiDAR are favoring novel approaches, also within a comparative perspective, based on the complementarity and symbiotic relation between urbanism and anthropogenic landscapes, see e.g., Chase and Chase (2016).
20 e.g., Polybius, Historiae, V.1.56 or Titus Livius, Historiae, I.12.
23 Gonzalez-Ruibal and Ruiz-Galvez (2016).
24 Naglak and Terrenato (2019).
25 Engel (1884).
suggested by Renato Peroni\textsuperscript{26} and Andrea Cardarelli\textsuperscript{27}, in their elaboration and definition of proto-urban societies and seems most promising.

Zuiderhoek’s book, these discussions and the rich literature of comparative studies on urbanism\textsuperscript{28} demonstrate that while the debate on what is an ancient city is still very much open and far from being resolved, it is still possible to identify some common traits and or common trajectories that characterize settlements and communities across a great variety of historical and/or chronological settings. However, much of the discussion of these themes, within historical and archaeological circles, has been on a discursive or qualitative level, therefore it is often difficult to harmonize the different models that have been applied to date into a consistent empirical and/or theoretical framework. A new approach to settlements throughout different contexts should now be within our grasp, however, thanks to the facilities that recent developments in IT offer us to model, analyse, and statistically test data. As suggested by Monica Smith “the capacities for human interaction in concentrated locations are exercised within a limited set of parameters”\textsuperscript{29}, that should be possible to study quantitatively. Zuiderhoek seems to be skeptical about these interdisciplinary and quantitative comparative approaches to urbanism and urbanization that “may eventually be able to arrive at some universal understanding of urbanism”\textsuperscript{30}. Differently I believe that qualitative discussion and comparative quantitative approaches are not alternative but complementary and it is still possible to keep details about cultural-historical specificity within wider comparative perspectives. In this sense Zuiderhoek underestimates a whole tradition of studies from the pioneering work by Louis Wirth\textsuperscript{31} to the more recent contributions by Michael Batty\textsuperscript{32}, both discussed and presented in the recent quantitative approach to Central European urbanism by Oliver Nakoinz\textsuperscript{33}.

The quantitative comparative approach presented in this Research Topic, allows us to connect recent developments in archaeological research with those in other disciplines, including economics, anthropology, sociology, and social ecology, not only enabling us to add historical depth to our models of urbanism, but also to connect understanding about cities in the past and present, offering opportunities to predict their evolution and improve policies in the future. Probably given my personal background and expertise, the collection is slightly biased toward Mediterranean cultures and classical civilizations, with a special focus on Italy, but probably this is not totally a bad thing since classical civilizations lay at the origin of Western culture, therefore understanding them better is also understanding ourselves a bit better, as long as we are aware of this potential bias and perspective.

Chapman et al., examines a particular form of early urbanism, in 4th millennium BC Ukraine, the so-called Trypillia Megasites, the largest known settlements of that time in Europe and possibly in the world. These sites are often viewed as failed experiments on the path to proper urbanism or proto-urban sites, and reveal few signs of hierarchical social stratification despite their large size; as such, they represent a challenge for the understanding of early processes of community formation and social integration. Chapman et al. consider inter-sites exchange and interaction and observes how Tryphillia Megasites’ subsistence stresses begin when site size exceeded the critical size of 35 ha. This tends to happen, especially in the passage between phase B1 and C1, when also a particular level of agglomeration and clustering can be noticed, suggesting that some form of buffering involving exchange of goods for food was in operation. In addition, by analyzing the layout and internal organization of these mega-sites, Chapman et al. suggest that they might also considered centers for religious agglomeration and processional rituals, which might be at the origin of their development and growth. This connection between the origin of urbanism and religion is also a novel perspective recently suggested by Jorg Rüpke, that applies very well also to other Mediterranean cultures including early Rome\textsuperscript{34}.

By using a wide range of data and by applying socio-material network analysis (community detection techniques) Mazzucato, in the Neolithic site of Catalhöyük, in Turkey, analyses households as nodes and investigate family and intra-community ties and relations. Here analysis reveal that by the later part of the Neolithic period, the houses network together with a low global density score, accounts for the highest centralization value. This configuration suggests a much less cohesive settlement in this period, where there is an increase of the central role of some buildings, together with a general contraction of material relations, which might indicate a more dispersed and less egalitarian social arrangement.

Households and intra-site analysis is also the focus of Cabaniss’s paper, which, by using the case study of Metaponto, introduces to the archaeological literature, the entropy estimating statistical bootstrap (EESB), a tool developed in information theory and computational social science by DeDeo et al. (2013). This tool is important, because provides a way to assess how representative a small dataset is of a parent population, categorized according to some useful typology, and therefore can be used to decide when small datasets can add further detail to our quantitative studies of archaeological settlements or when they need to be rejected as too small. As emphasized by Cabaniss, “this will allow building larger urban datasets that are empirically grounded in the specific evidence for each community, facilitating the work of research programs such as urban scaling.”

Similarly to Mazzucato’s paper, The Davis model of community detection has also been used by Donnellan to explore community dynamics at an emerging indigenous urban site in Southern Italy, which showed signs of intense contact...
both with Etruscan and Greek communities (900–600 BC). By using two-mode model networks between burials and grave goods objects, she calculated different indexes such as network cohesion and centrality measures. Network cohesion showed expanding and contracting, suggesting probably the existence of tension and a tight control of funerary behavior from the community. In addition, the study of centrality of selected nodes suggested that an increase in crop storage has played a significant role in the development of state power and the urbanization process at Pontecagnano.

Stoddart et al. authors combine the use of rank size and indices of centralization at the regional and local level, by examining the large places and the supporting rural settlements, by using survey data from many projects conducted in central Italy since the second half of the nineteenth centuries and especially in the second half of the twentieth century, and recently made available also thanks to innovative open-access digital platforms. This paper examines the power dynamics as indicated by settlement organization in Etruria during the first half of the first Millennium BC and identifies areas of greater centralization and some areas of vacuum of power. The overall picture is similar to my own recent study, and identifies the main distinction between Etruria and Latium in the difference balance of powers and suggests that an “Etruscan empire” was unlikely because of the specific heterogeneous structure of Etruscan communities and settlements organization.

Nakoinz et al. use a particular type of artifact, fibulae which are a garment of dress and a personal ornament common in Iron Age Mediterranean and Continental Europe societies, to build “middle class” networks among German princely seats and translate Christaller relative centrality into network centrality. By adopting and combining concepts from different tradition of studies, such as urbanity, centrality, interaction, and connectivity, they offer a case studies and develop a methodology that allow to combine social and geographical networks in a novel and promising way.

Mandich’s article investigates the urban expansion and economic development of ancient Rome through the application of models and theories originally designed for the study of contemporary cities. While the growth of ancient settlements is often difficult to track and analyse, archaeologically observable changes in land use can be read and interpreted as a function of broader economic oscillations over the longue durée. In particular, Mandich shows, how specific patterns of urban expansion identified in modern cities also existed in ancient Rome.

Fletcher’s paper compares urban settings in different region and chronological settings and defines different types of urbanism according to different density patterns. In particular, he compares modern industrial cities to pre-industrial agrarian societies, and he identifies two different types of low-density: large low-density settlements on a grand scale, in the range of 200–1,000 ha (Greater Angkor, Mayan cities, Tryphillia Megasites etc.) and low density settlements of a lesser scale, between 15 and 20 ha sometime even 80–90 ha (Great Zimbabwe, European oppida etc.). While the first ones seem not to have long-term trajectories, the smaller but well-connected small scale, low-density settlement seems to have longer trajectories and sometime, eventually develop into the industrial modern cities.

A common thread of all these papers seem to be the recognition that ultimately “urbanity” is mainly “connectivity” and probably within the traditional dichotomy between “hierarchy” and “heterarchies” lies the clue for the development of “urban” complexity. Translated in other terms, Smith and Lobo argue that the variability among cities, can be in essence reduced to two basic types of urban types: political (most ancient cities) and economic (most modern cities). However, both these types can be reconciled in an ultimate model of spaced urban environment, which again is based on connectivity: cities as settings for “energized crowding.” As Smith and Lobo suggest, processes of interaction generate both economic and political growth, and they ultimately produce and influence the built forms and social characteristics of all cities.

This model may help scholars distinguish the unique from the universal traits of cities today and in the past. In his second paper of this special Research Topic, Bettencourt and Lobo, this time with Bettencourt, discuss quantitative comparisons based on a few simple variables across settlements to analyse how different places and peoples dealt with general problems of any society. These include demographic change, the organization of built spaces, the intensity and size of socioeconomic networks and the processes underlying technological change and economic growth. As reminded by Bettencourt and Lobo, the historical record contains a much more varied and more independent set of experiences than contemporary urbanization, it has a unique power for illuminating present puzzles of human development and testing emergent urban theory.

In his paper, Ortman, follows up on this argument. Past developing urban contexts provide a diachronic laboratory to assess different socio-economic factors to determine how and why urban environments came into being, developed, flourished, and eventually collapsed (or not). However, as emphasized by Ortman, often lessons from the past can be hindered by the fact that they remain anchored to a very peculiar and specific chronological and geographical context. By partially going back to the unfulfilled potential of some of the New Archaeology aspirations, and by adopting quantitative comparative perspective, such as settlement scaling theory, we can overcome these limitations, and archaeology could assume a “new kind of relevance” that goes beyond rhetoric declamations. We hope this collection of paper, presenting both case studies and theoretical essays has offered some material and opportunity for discussion in this direction.

Probably what has been left partially implicit in this collection, is the experiences of people who live and work in these urban environments, their well-being which ultimately is also a measure of economic growth: where and how people live, eat, travel, and interact? How does people’s life change as communities become increasingly urban? What are the health

---

35Fulminante (forthcoming).
36See also pioneering observations in this sense already in Guidi (1985).
37On the challenging opportunity of comparing and combining social and geographical networks see also the recent volume by Dawson and Iacono (2020).
differences between urban and rural populations and/or people of different social status? But these are questions for another Research Topic.

**AUTHOR CONTRIBUTIONS**

The author confirms being the sole contributor of this work and has approved it for publication.

**ACKNOWLEDGMENTS**

This Special Research Topic stemmed from a collaboration and discussion originated at the “Urbanization without Growth in the Ancient World” Conference, held at the British School at Rome, 18th–21st July 2017, and organized by the Social Reactor Project (Colorado). The editorial work is partially supported by this project, and by grants held by Francesca Fulminante at the Institute of Advanced Studies, at Durham (Epiphany Term, 2018) and at Kiel (DIG Visiting Fellowship Michaelmas Term 2018) Publication of the research presented in the papers has been supported by an OPENAIRE funding linked to the Marie Curie Project Past-people-nets (628818), conducted by Francesca Fulminante (2014–2016) and by a grant from the James S. McDonnell Foundation (#220020438). I would like to thank the other editors of the Research Topic for organizing such a stimulating conference in Rome a few years ago and for editing and contributing to the Research Topic and the other authors for their great work.

**REFERENCES**


Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Fulminante. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Socio-Material Archaeological Networks at Çatalhöyük a Community Detection Approach

Camilla Mazzucato*

Department of Anthropology, Stanford University, Stanford, CA, United States

Vast in scale and densely inhabited, Late Neolithic Near Eastern megasites have been variously considered in relation to urbanity. Often viewed as failed experiments on the path to proper urbanism or proto-urban sites, these settlements reveal few signs of hierarchical social stratification despite their large size; as such, they represent a challenge for the understanding of early processes of community formation and social integration. Drawing upon a wide range of data and using socio-material network analysis as a methodological tool, this paper explores the way the late Neolithic site of Çatalhöyük was organized internally and specifically the way individual houses were embedded in the wider social fabric of the site. This study sheds light on the nature of the networks of social engagement and affiliation that emerge in the Holocene within large early agricultural communities and the way such networks were manifested.

Keywords: Neolithic, Çatalhöyük, Near East, network analysis, community detection, cities

INTRODUCTION

What is a city? When can we talk about an “urban way of life”? And where do cities come from? The debate surrounding “urbanism” has a long and extremely complex history. The issue of what constitutes a city is an inexhaustible topic that is endlessly readdressed and reexamined through a variety of lenses. Investigating urbanism, especially early urbanism, implies having to address an array of crucially interwoven issues that encompass the processes of sedentarization, the intensification of social complexity, the construction of new forms of personal and group identities, negotiation and conflict resolution, the changing nature of the relationship with the natural and the material environment and the shift to a way of life increasingly reliant on agriculture and animal husbandry. The emergence of forms of permanent residence in population dense large-scale settlements during the Neolithic period involved deep transformations of the hunter-gatherer ethos and way of life, so radical that it has been suggested that the process lead to essential cognitive transformations (see Benz and Bauer, 2013; Watkins, 2013; Sterelny and Watkins, 2015; Benz, 2017). Processes of group integration, of community construction and establishment of mechanisms of social regulation started very early in the Epipaleolithic and Early Neolithic (e.g., Kuijt, 2000; Kuijt and Goring-Morris, 2002; Hodder, 2005, 2014a, 2018; Goring-Morris and Belfer-Cohen, 2010, 2011; Belfer-Cohen and Goring-Morris, 2011, 2013, 2017; Benz et al., 2013, 2017; Benz, 2016; Finlayson and Makarewicz, 2017), but it is in the PPNB, with the emergence of large clustered agglomerations (megasites), that such processes “scale-up” to an extent and an intensity that could be interpreted as “almost urban” (Mazzucato, 2016). Jordan is home to the earliest appearance of these settlement types, but others megasites emerged across the late Neolithic Near

Within the context of the prehistory of the Near East, the debate surrounding “the city” has been and is still strongly influenced by an enduring quest for the origin of the urban phenomenon (see Bienert, 2001; Gebel, 2002; Rollefson, 2004; Yoffee, 2005; Ben-Shlomo and Garfinkel, 2009). This search primarily developed within a predominant evolutionary framework that views the “city” as emerging out of a linear progression that starts with small settlements and ends up around the 4th millennium with the large southern Mesopotamian Bronze Age urban agglomerations (Ben-Shlomo and Garfinkel, 2009; Gaydarska, 2016, 2017; Ur, 2017). These large and socially stratified settlements located in what is now Iraq typify the canonical form of the early city. This process of spatial and demographic expansion and the parallel increase in nucleation are envisioned as the prelude to state formation and are intimately related to the gradual emergence of hierarchically arranged social forms (see Yoffee, 2005). Childe was the first to observe the flourishing of Mesopotamian cities and to describe it as an abrupt change—an “urban revolution” (Childe, 1950)—a threshold in the evolution of human social forms that brought about a number of changes and features (e.g., the invention of writing and the development of centralized administrations) (Childe, 1950). We know now that the Mesopotamian landscape was punctuated by a number of vast and dense population agglomerations that predate the appearance of the 4th millennium “cities” by a millennium, for example, Tell Brak (Wengrow, 2010; Ur, 2014, 2016, 2017). Sites like Tell Brak are usually defined as “proto-urban” and considered part of those “successful” experiments that flourished through the adoption of social and political institutions to mitigate inter-community conflicts, ultimately anticipating the advent of “proper” cities such as Uruk in southern Mesopotamia (Oates et al., 2007; Ur, 2017). Within this framework, population density, settlement nucleation and the origin of centralized institutions and stratified societies developed together and are deeply correlated. This generally assumed strict correlation is challenged, however, by a series of finds and settlement types (Wengrow, 2015). It is worth noticing here, that early forms of centralized administration, recording and storage management are attested from the 6th millennium BCE within small agriculture villages like Tell Sabi Abyad in Syria (Akkermans and Verhoeven, 1995; Wengrow, 2010; Akkermans, 2014). Additionally, the Neolithic megasites complicate this narrative, since they are both population dense, inhabited for a long time and lack any clear sign of centralized systems of administration or hierarchical arrangements of power. Neolithic megasites, such as Çatalhöyük, have been variously defined as “severe anomalies” (Fletcher, 1995, p. 189) or “dead ends” (Ben-Shlomo and Garfinkel, 2009, p. 203) on the way to true urbanity; alternatively, they have been viewed as proto-urban sites or “proto-cities”: early “experiments” with social nucleation that didn’t continue any further (Ur, 2017, p. 140). Instead, these settlements and finds seem to point to a much more flexible and complex scenario of multiple trajectories and experiences that can be hardly restricted within linear and univocal narratives and that suggest the need for a focused contextual approach and a bottom-up perspective that rather of trying to restrict the different settlement forms and practices within normative categories is concerned with the way these sites were internally organized, on which socio-material practices formed their fabric and how they changed through time and space (Hodder, 2005; Asouti, 2006; Düring, 2007a,b, 2013; Wengrow, 2015; Mazzucato, 2016; Der and Issavi, 2017).

This paper is part of an ongoing program of research that explores mechanisms of social integration and group formation within the site of Çatalhöyük. Using a socio-material network approach and a community detection method, it is focused on identifying and disentangling the dynamics of interconnectivity and patterns of affiliation and cooperation between buildings and how these reflect the material choices and the spatial organization at the site. Furthermore, it aims to highlight social units larger than the single house (e.g., neighborhoods, corporate groups or sodalities) and to inspect the way they were embedded within the wider site. This study contributes to the debate regarding the way megasites were internally organized and their forms of social integration furthermore, it contributes to the discussion regarding forms of Neolithic corporate identities (see Hodder and Pels, 2010; Hodder, 2014b; Bogaard, 2015; Benz et al., 2017; Kuijt, 2018). Socio-material network methods are used as the methodological tool for investigating these issues because they provide the opportunity to consider connectivity and dependencies between units of analysis in a synthetic way that incorporates different material classes.

THE ANATOMY OF ÇATALHÖÜK

Çatalhöyük is typically considered one of the larger Near Eastern Neolithic settlement sites (13.5 ha) (Figure 1). First discovered and excavated by James Mellaart between 1961 and 1965, it has been investigated by the Çatalhöyük Research Project (ÇRP) under the direction of Ian Hodder since 1993.

The Neolithic occupation at Çatalhöyük extends for more than a millennium (7100–6000 cal. BC) (Bayliss et al., 2015) and is characterized by sequences of mudbrick buildings constructed one on top of the other and, at intervals, separated by external spaces such as middens and penning yards. Çatalhöyük shows no signs of deliberate planning; instead, the site seems to have developed an organic, modular arrangement through the repetition of similar structures. Mudbrick buildings were densely packed in a close-knit fabric that did not allow for the presence of roads or significant open areas. As with other Neolithic megasites (e.g., Basta, Ba’ ja, Aşikli Höyük), Çatalhöyük’s houses formed the center of domestic and ritual life and were characterized by a highly regular repetition of the same elements, while at the same time revealing smaller idiosyncrasies (in terms of size, layout, material culture) that resulted in each house being slightly different and independent, and having a specific identity (Asouti, 2006; Hodder, 2014a; Hodder and Farid, 2014).
FIGURE 1 | Çatalhöyük East Mound plan. Main excavation areas highlighted in gray. (Plan: Camilla Mazzucato. Used by permission of the Çatalhöyük Research Project).
Despite being constructed very close to each other, houses at Çatalhöyük are, throughout the entire sequence, easy to distinguish spatially. Shared party walls are very rare and mainly confined to the early levels; instead, houses are typically delimited by independent walls (Hodder, 2005; Düring, 2007a,b). These independent rectilinear mudbrick walls usually delimit a central living area and smaller side rooms used for storage or food processing. Except for some houses in the early occupation of the site, the living area is typically divided into a northern sector where burials, wall paintings, installations and, in general, symbolic features were located, and a southern sector where fire installations (e.g., ovens and hearths) are located and food preparation and domestic activities took place (Hodder, 2005, 2013, 2014a; Hodder and Farid, 2014). Throughout the entire sequence, houses at the site are the focus of burial activity, symbolic elaboration, small-scale food processing, consumption and small-scale production activity (Bogaard et al., 2009; Bains et al., 2013; Carter and Milic, 2013; Demirerghi et al., 2014). While the houses at Çatalhöyük were clearly durable spaces in which the concerns of the site’s inhabitants for memory and time depth are revealed through repetitive practices, they were also dynamic and continually evolving structures. Houses at Çatalhöyük were repeatedly modified over the course of their occupation. Wall installations and paintings were ephemeral features that changed frequently, together with many of the houses’ internal structural elements (Hodder, 2005, 2013, 2014a; Düring, 2007a,b; Hodder and Farid, 2014).

Here some terminological clarifications should be made before carrying on. In this study the term building and house will be used interchangeably, and they refer to the above described spatial unit whose modular repetition characterizes Çatalhöyük’s clustered configuration. The term household will point, instead, to a more flexible entity, a unit of people or a “social group” that may or may not coincide with the architecturally defined house, which is defined by cooperation within the social, economic (sharing of economic resources) and ideological “sphere” (Düring, 2007a,b; Souvatzi, 2007, 2008, 2012). In addition to the small-scale nature of consumption and production observed, and the lack of evidence for inequalities between houses, the modular nature of the buildings and their remarkable uniformity suggests the possibility that buildings at Çatalhöyük represent independent units, i.e., “discrete household residences” (Düring, 2007b, p. 163).

While no clear evidence for social stratification has been observed archaeologically at Çatalhöyük, horizontal differences and variations between buildings have been detected. These variations were first explained by Mellaart through the lens of a stratified society. He characterized some of the most elaborate houses he excavated as “shrines,” and the areas where these “shrines” tended to cluster as “priestly quarters” (Mellaart, 1967). Since the beginning of the Hodder project, however, interpretations of the differences between buildings and the social geography of the site have diverged from those of Mellaart. Hodder and Pels (2010) see the distinction between ordinary houses through the lens of Lévi-Strauss’ concept of house societies (1979); here they make a distinction between standard houses and “history houses” and suggest the existence of complex series of dependencies between them (Hodder and Pels, 2010). History houses suggest a process of differentiation involving the accumulation of social memory that is materialized in symbolic elaborations such as wall paintings and plastered animal installations that marked these buildings as important focal points for community subgroups (Düring, 2007b; Hodder and Pels, 2010). Not dissimilarly, Düring and Marciniak (2006) highlight the role of “social associations” beyond the house as a principle of social organization at Çatalhöyük and at other Anatolian Neolithic sites (e.g., Aşıklık Höyük, Canhasan III). Houses are viewed as embedded in clustered neighborhoods, possibly as satellites of particular focal buildings (Düring and Marciniak, 2006). Within this framework, spatial proximity represents a very important organizing principle of the social fabric that can be observed as it changed through time. A similar supra-household organization is reflected in architectural clustering and forms of ritual affiliations at other megasites, for example at A’in Ghazal (Rolleson, 2015). Less concerned with spatial location as a factor in social organization is the interpretation of Çatalhöyük’s social arrangement in sort of “social mosaic” (Hodder, 2014b, p. 151) formed by “flexible networks” (Mills, 2014, p. 179) of cross-cutting social groupings (e.g., religious sodalities) that overlapped to form a “tightly-knit and highly successful society” (Hodder, 2014b, p. 167). Mills (2014) suggests that history houses acquire prestige and power through networks of affiliations and relations with other houses. Both Hodder and Mills envision a social structure comprised of autonomous houses situated within cross-cutting series of relationships and dependencies with other houses. In a similar vein, Bogaard and co-workers refer to forms of cooperative farming at Çatalhöyük and at other PPN-PN sites (Bogaard, 2015, 2017; Bogaard et al., 2017). Kuijt (2018) has recently suggested that the social geography of Çatalhöyük can best be characterized as an expression of households as multi-family houses. Applying Lévi-Strauss’ concept of house society, Kuijt attributes specific function to particular houses (e.g., places for burying the dead or places for symbolic elaboration) used by an extended/multi-family household made up of different families/components kept together by affiliation/membership spatially spread in clusters of structures.

Recent research at Çatalhöyük has highlighted a clear shift in social arrangements between an early phase (pre XII level–Levels South O/North G–before 6500 B.C.) characterized by a complex entanglement of dependencies between social groups and a late phase during which it appears that single households emerge as more independent units and the social organization of the community appears more fragmented and dispersed (Levels South P/North H–TP–after 6500 B.C.). The codified, collective social practices observed in the early levels are abandoned and a more autonomous, household-based expression of community emerges (Hodder, 2013, 2014b; Marciniak, 2015; Marciniak et al., 2015). These developments should be viewed within broader regional and Anatolian trends, especially the overall tendency toward household autonomy and production intensification (Flannery, 1972, 2002; Boyd, 2005; Düring, 2006, 2007a; Kuijt, 2008; Hodder, 2013). Despite these observed temporal changes, however, the population seems to have maintained an egalitarian
ethos until the abandonment of the East Mound at the end of the Neolithic (~6000 cal. BC). For example, no major differentiation in terms of storage capacity or intensity of production within houses has been observed.

Relational models, such as the ones employed in the current study, appear to well represent the set of horizontal ties that structured society at the site, making Çatalhöyük and other megasites an ideal place to explore the potentialities and limits of socio-material archaeological network approaches as a means of answering archaeological questions.

TYING THE KNOT: SOCIO-MATERIAL SIMILARITY NETWORKS AT ÇATALHÖÜK

Socio-material archaeological network methods fall within the discipline of Network Science (Brandes et al., 2013). Brughmans (2010. p. 277) defines them as a means of “detecting and interpreting patterns of relationships between subjects of research interest.” They are based on the premise that a series of archaeologically observed phenomena can be abstracted as systems that “exhibit an interdependent organization” and that, within these systems, sets of associations between elements can be conceptualized as links that form network models (Brandes et al., 2013. p. 3). Network models are abstractions formed of individual elements or entities represented as nodes (or vertices) and connected by lines (or links/ties) that represent some form of relation between nodes. As such, entities are analyzed in terms of their embeddedness within sets of relations; they are therefore never autonomous but always dependent on the behavior of all the other entities and their connections (Borgatti et al., 2009; Marin and Wellman, 2011). This type of approach is built on the observation that, instead of focusing on entities in isolation, analyzing the connections or the set of relations between them provides a much deeper understanding of the dynamics of specific phenomena (Borgatti et al., 2009; Knappett, 2011; Brughmans, 2013; Mol, 2014; Collar et al., 2015; Brughmans et al., 2016).

At Çatalhöyük the entities/nodes in the network are represented by individual buildings whose relationships to other buildings are established based on the repeating of specific material features. Socio-material networks are, therefore, traced through similarities between material culture and through the co-occurrence of specific features that are used as the way of constructing links among entities. Individual buildings at Çatalhöyük have been chosen as the major system entities (nodes) because they are the key component of the social fabric and, despite differences and changes through time, they represent the main and enduring principle of social organization (Tringham, 2000; Hodder and Pels, 2010; Souvatzi, 2012; Düring, 2013; Baird et al., 2016). The house—with its location in the landscape and the materials associated with it—offers the necessary link for network modeling. Thus, it is the “materiality, spatiality, historicity, and specificity” of buildings that make them appropriate analytical units at the site (Souvatzi, 2012. p. 4).

The material features that are used to establish relationships between buildings and to infer their shared affiliation, have been

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Objects/practices.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
<td><strong>Features (presence/absence)</strong></td>
</tr>
<tr>
<td><strong>Architecture</strong></td>
<td>Shape of buildings</td>
</tr>
<tr>
<td></td>
<td>Location of highest platform</td>
</tr>
<tr>
<td></td>
<td>Location of platform with the highest amount of burials</td>
</tr>
<tr>
<td></td>
<td>Location of benches</td>
</tr>
<tr>
<td></td>
<td>Location of posts</td>
</tr>
<tr>
<td></td>
<td>Location and typology of oven</td>
</tr>
<tr>
<td></td>
<td>Location and type of bins</td>
</tr>
<tr>
<td></td>
<td>Separation between oven and hearth</td>
</tr>
<tr>
<td></td>
<td>Presence of screen/kerb in the main room</td>
</tr>
<tr>
<td></td>
<td>Elaboration on the West wall</td>
</tr>
<tr>
<td></td>
<td>Burnt status</td>
</tr>
<tr>
<td></td>
<td>South-West corner</td>
</tr>
<tr>
<td><strong>Burials</strong></td>
<td>Isolated heads in burials</td>
</tr>
<tr>
<td></td>
<td>Headless bodies</td>
</tr>
<tr>
<td></td>
<td>Wooden plank</td>
</tr>
<tr>
<td></td>
<td>Rodents</td>
</tr>
<tr>
<td></td>
<td>Red ochre</td>
</tr>
<tr>
<td></td>
<td>Cinnabar</td>
</tr>
<tr>
<td></td>
<td>Heads in tertiary contexts</td>
</tr>
<tr>
<td></td>
<td>Foundation burials</td>
</tr>
<tr>
<td></td>
<td>Presence of secondary burials</td>
</tr>
<tr>
<td></td>
<td>Shells in burials, types</td>
</tr>
<tr>
<td><strong>Faunal remains</strong></td>
<td>Birds</td>
</tr>
<tr>
<td></td>
<td>Deer antlers</td>
</tr>
<tr>
<td></td>
<td>Specific parts of animals (e.g., wolf paw)</td>
</tr>
<tr>
<td></td>
<td>Bone fish hook</td>
</tr>
<tr>
<td></td>
<td>Bone harpoon</td>
</tr>
<tr>
<td></td>
<td>Bone pendants</td>
</tr>
<tr>
<td><strong>Botanical remains</strong></td>
<td>Lentils</td>
</tr>
<tr>
<td></td>
<td>Peas</td>
</tr>
<tr>
<td></td>
<td>2-raw naked barley</td>
</tr>
<tr>
<td></td>
<td>Emmer</td>
</tr>
<tr>
<td></td>
<td>Acorns</td>
</tr>
<tr>
<td></td>
<td>Almonds</td>
</tr>
<tr>
<td></td>
<td>Crucifix</td>
</tr>
<tr>
<td><strong>Obsidian</strong></td>
<td>Obsidian hoards</td>
</tr>
<tr>
<td></td>
<td>Bifaces in hoards</td>
</tr>
<tr>
<td></td>
<td>Projectiles at abandonment</td>
</tr>
<tr>
<td></td>
<td>Opposed platform blades</td>
</tr>
<tr>
<td></td>
<td>Opposed platform blade core</td>
</tr>
<tr>
<td></td>
<td>Pressure blade</td>
</tr>
<tr>
<td></td>
<td>Pressure blade/leaver</td>
</tr>
<tr>
<td></td>
<td>Pressure blade long</td>
</tr>
<tr>
<td></td>
<td>Mirrors</td>
</tr>
<tr>
<td></td>
<td>Evidence of manufacture</td>
</tr>
<tr>
<td><strong>Chert</strong></td>
<td>Flint dagger</td>
</tr>
<tr>
<td></td>
<td>Evidence of manufacture</td>
</tr>
<tr>
<td></td>
<td>Pressure-Lever technique</td>
</tr>
<tr>
<td></td>
<td>Percussion technique</td>
</tr>
<tr>
<td></td>
<td>Percussion-Indirect technique</td>
</tr>
<tr>
<td></td>
<td>Knife pressure technique</td>
</tr>
<tr>
<td></td>
<td>Knife indirect percussion technique</td>
</tr>
<tr>
<td></td>
<td>Knife pressure crutch technology</td>
</tr>
<tr>
<td></td>
<td>Knife pressure</td>
</tr>
<tr>
<td><strong>Personal adornment</strong></td>
<td>Evidence of manufacture of beads</td>
</tr>
<tr>
<td></td>
<td>Butterfly beads</td>
</tr>
<tr>
<td></td>
<td>Decorated boar tusk collar</td>
</tr>
<tr>
<td></td>
<td>Beads shells</td>
</tr>
<tr>
<td><strong>Wood</strong></td>
<td>Type</td>
</tr>
<tr>
<td><strong>Bricks</strong></td>
<td>Brick types</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 1 | Continued

<table>
<thead>
<tr>
<th>Category</th>
<th>Features (presence/absence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paintings</td>
<td>Paint type (net, honeycomb, zigzag, bricks, headprints, anthropomorphic, vulture, bull, leopard, rhomb, checker board, triangles)</td>
</tr>
<tr>
<td>Installations</td>
<td>Installation type</td>
</tr>
<tr>
<td>Groundstone</td>
<td>Maceheads</td>
</tr>
<tr>
<td></td>
<td>Decorated querns</td>
</tr>
<tr>
<td></td>
<td>Quern</td>
</tr>
<tr>
<td>Pottery</td>
<td>Pottery placement</td>
</tr>
<tr>
<td>Figurines</td>
<td>Anthropomorphic figurine in cluster of objects</td>
</tr>
<tr>
<td>Closure practices</td>
<td>Placed dismantled installations</td>
</tr>
<tr>
<td></td>
<td>Dumped feasting remains</td>
</tr>
<tr>
<td></td>
<td>Placed items ground stone</td>
</tr>
<tr>
<td></td>
<td>Spread of cattle bones/abandonment</td>
</tr>
<tr>
<td></td>
<td>Missing closing deposits</td>
</tr>
<tr>
<td></td>
<td>Oven closure (and bin) preserved and filled</td>
</tr>
<tr>
<td></td>
<td>Spread of botanical remains</td>
</tr>
<tr>
<td></td>
<td>Closure deposit in oven</td>
</tr>
</tbody>
</table>

.selected from the full range of archaeological data recorded at Çatalhöyük between 1993 and 2017 (Table 1). These data should be understood as the “material traces” of the processes of producing “social relations” in everyday life (Joyce, 2015, p. 185) and as footprints of the “choices in the way people engaged with objects” and of the act of reproducing social relations (Mills, 2016, p. 247). They represent situated practices indicating communalities that link households to form communities within the overall site (Pollock, 2006). Using all the material classes together has the advantage of considering all materials in a sort of “flat,” non-hierarchical way that avoids dichotomizing the dataset between ritual and domestic or private and public. At Çatalhöyük these differences (e.g., between domestic and ritual) are always difficult to disentangle and various elements are intertwined in such a manner that separating or differentiating between them is impossible (Hodder, 2014a).

This approach is based on the assumption that the higher the level of similarity between buildings, the higher the probability that they were somehow affiliated; this assumption forms the basis of previous archaeological network analyses (see Östborn and Gerding, 2014 for a review and Mills et al., 2018; Giomi and Peeples, 2019 for recent applications). Furthermore, this approach acknowledges that the complexity of Çatalhöyük’s history can be grasped only by combining the widest possible number of datasets, even if they are not obviously related, and that a relational and contextual process of “assemblage” of strands of evidence is needed to produce robust interpretations (Hodder, 2014a, 2015).

A large part of this dataset is the architectural elements that form the structure of houses (e.g., location of ovens and benches or location of highest platforms) as well as the burial assemblages within them (e.g., the presence of secondary burial types or pigments or of specific shell species). Additionally, a vast range of other object types and data categories that form the assemblages of buildings have been used to create connections (Table 1). Some material features have been selected because they represent a sort of variation from the normative setting of houses, for example ovens located in the southern part of buildings is not recorded in this network dataset because this represents the usual location of fire installations at Çatalhöyük. In contrast, the presence of ovens located in the northern sector of buildings, or ovens that shift to the southwest corner, have been recorded as they provide more information on variations in practice on site.

It should be highlighted here that this approach acknowledges that network models created using archeological data are different from straightforward social networks in that they aim to reconstruct social dynamics, but they can only be created using material culture as a proxy for social relations (Knappett, 2011, 2016; Brandes et al., 2013; Hodder and Mol, 2015; Peeples et al., 2016; Van Oyen, 2016). Strictly speaking, these networks are neither social nor material but “socio-material” in nature (Knappett, 2011). In this method, society and material culture are conceived as interdependent systems.

For the purpose of this research I will use two different types of networks: 2-mode networks and 1-mode weighted networks of both buildings and artifacts (Figure 2). All the buildings excavated by the ÇRP are arranged within three main temporal groupings or periods (early, middle, late) (Table 2). These “macro-phases” are obtained by collapsing several stratigraphic levels at Çatalhöyük into larger chronological periods in order to generate networks that are chronologically consistent and populated by an adequate number of nodes for analysis.

For each macro-phase the presence/absence of specific objects or practice are recorded as a nodelist (nodelist2 – UCINET) (Figure 2A) which is then displayed as a 2-mode network (Figure 2B). 2-mode or affiliation networks are made of two different type of nodes, in this case buildings and “objects,” and of ties that link directly only the different node types (Borgatti and Everett, 1997; Prell, 2012; Borgatti et al., 2013). The 2-mode network can be then projected as a weighted 1-mode network of buildings (Figure 2C) or “objects” (Figure 2D); for the purpose of this article I will focus on both the weighted network of buildings and of objects. The 1-mode network of buildings links individual buildings (nodes) through weighted ties; the weight of ties is determined by the number of objects shared by dyads of buildings. The projected 1-mode network of “objects” links the artifacts/practices that are recorded as co-present in the same building. Therefore, the weight of links represents in these networks the amount of buildings shared by dyads of artifacts. These extracted 1-mode networks are proper archaeology similarity networks (ASN) as described by Prignano et al., “spatial networks derived as the one-mode version (projection) of (weighted) bipartite networks” (Prignano et al., 2017, p. 5), even if, for this study 1-mode networks are the result of the projection binary bipartite networks.

Two-mode networks are largely used for this study as a useful tool for better observing the way objects bring building together
FIGURE 2 | Relevant objects and buildings are recorded and arranged in a node list format (UCINET) (A) and displayed as a 2-mode network (B). 2-mode networks or affiliated networks are made of 2 type of nodes in this case of buildings and objects that are linked (affiliated) to one another. The 2-mode network can be converted in weighted 1-mode networks of buildings (C) or objects (D). The weighted 1-mode network of buildings links the building (nodes) that share the same objects. The weight of links represents the amount of object that link each couple of buildings. The 1-mode network of objects on the contrary links the objects that are co-present in the same building and the weight of the link correspond to the number of buildings shared by couple of objects.
in networks; they are a more transparent way of investigating the relationship between “people” and artifact types within “mutually constitutive networks” (Knappett, 2011; Mol, 2014, p. 89).

Community detection analysis has been performed only on 1-mode weighted networks of buildings to determine clusters of buildings that show a tighter material connectivity between themselves compared to buildings that belong to other groups. Generally, networks are not binarized since weights of ties provide the research with extremely important information regarding the intensity or possibly duration of connection (Peeples and Roberts, 2013).

## DETECTING COMMUNITIES IN NETWORKS

Social networks tend to have a highly inhomogeneous structure characterized by an arrangement that appears to naturally subdivide into areas of tightly interconnected nodes and, subsequently, of high concentration of edges. These groups of nodes, that are very dense within themselves but display exigous connection with other groups are called communities, clusters or modules (Newman, 2003a, 2006; Fortunato, 2010). Therefore, within communities, nodes display a higher intensity of interaction and as such, they have a much higher probability of forming links with their neighborhood nodes than with vertices outside of their community (Newman and Girvan, 2004; Newman, 2006, 2016; Fortunato and Hric, 2016) A wide number of methods and algorithms are now available and routinely used to expose community structure of networks and to perform community detection in different types of networks (see Lancichinetti and Fortunato, 2009, 2011; Fortunato, 2010; Fortunato and Hric, 2016; Yang et al., 2016 for a complete review of methods and critical analysis). These methods are based on different criteria and different ideas of what a community is. Among this variety of different tools modularity-based methods and algorithms that point to its optimization, are very popular and effective approaches to the problem of identifying communities in networks of modularity (Girvan and Newman, 2001; Newman, 2006, 2016). They are based on the quantification of modularity values and its consequent maximization (Newman and Girvan, 2004; Newman, 2006; Brandes et al., 2007; Fortunato, 2010). Modularity-based approaches are based on the idea that the best way to partition a network is to compare the structure and density of edges of an observed network with the density of edges expected if they formed purely by chance (Newman, 2006, p. 8578; Newman and Girvan, 2004). Modularity values capture a quantitation of such observation and they essentially measure the “number of edges that fall in within groups minus the expected number in an equivalent network if edges were placed at random” (Newman, 2006, p. 8578). Therefore, the modularity approach implies that the community structure of a network is always defined in comparison to a random similar network (Fortunato, 2010). Modularity is, additionally, a “quality function” that measures the “goodness” of network partitions and positive and “preferably” high scores of modularity are considered good indicators of sound and significant community structure (Newman, 2006, p. 8578; Fortunato, 2010). The optimization of this value is therefore the objective of a wide number of algorithms that seek to approximate to the maximum value of modularity possible (Q_max) as a way of decomposing the structure of a network. Modularity optimization algorithms are by far the widely used methods for community detection and they have been tested and extended to be used with weighted networks (Brandes et al., 2007; Lancichinetti and Fortunato, 2009; Yang et al., 2016). The method that has been proved to perform extremely well in comparison to both other modularity maximization methods (e.g., simulating annealing or Girvan and Newman) and algorithms based on different approaches (e.g., spectral algorithm, or Markov cluster algorithm) is the Louvain or multilevel algorithm (Blondel et al., 2008; Yang et al., 2016). This is a “greedy” optimization approach that achieves community modularity maximum in two steps: first, it takes into account all the nodes in the network and then, in the second step of the process, it utilizes the partitions obtained in the first phase to obtain a final maximum modularity score (Blondel et al., 2008). Due to this multilayered process and ability of capturing modularity at different levels of resolution this method is regarded as one of the most flexible methods of network decomposition. The layers of community detection are all to be regarded as meaningful partitions and allow for a multi-resolution observation of the network structure (Blondel et al., 2008).

Furthermore it has been demonstrated, through comparison with other detection methods using specific benchmark

---

**TABLE 2** | Çatalhöyük phases/levels and macro-phases used as chronological frame (early, middle, late).

<table>
<thead>
<tr>
<th>Temporal groupings of levels</th>
<th>South</th>
<th>North</th>
<th>Cal BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final</td>
<td>TPO-R and TPC</td>
<td>GDN</td>
<td>6300–6500 BC</td>
</tr>
<tr>
<td></td>
<td>Trenches 1 and 2</td>
<td>North H, I, J and IST</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(B109 and 115)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late</td>
<td>South T, TPN, TPC</td>
<td>South O</td>
<td>6600–6300 BC</td>
</tr>
<tr>
<td></td>
<td>B110 and B150</td>
<td>North F, G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South S, TPN, TPC</td>
<td>South O</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B150 and B122</td>
<td>South N</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South R</td>
<td>South M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South Q</td>
<td>South K</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South P</td>
<td>South I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South O</td>
<td>South L</td>
<td>7100–6700 BC</td>
</tr>
<tr>
<td>Middle</td>
<td>South O</td>
<td>South K</td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td>South L</td>
<td>South J</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South H</td>
<td>South I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South G</td>
<td>South H</td>
<td></td>
</tr>
</tbody>
</table>

*Mellaart level system was substantially reframed by the CPR after the 2008 season using the finer stratigraphic details gained during the Hodder years and the most recent understanding of the site’s material culture (Farid, 2014a).*
networks, that the Louvain algorithm, even with some problems with resolution, performs very well with networks that have few nodes, such as that of Çatalhöyük, achieving both accuracy and a fast computing rate (Fortunato and Barthelemy, 2007; Blondel et al., 2008; Yang et al., 2016).

Together with the $Q_{\text{max}}$ value the Louvain algorithm, like other community detection methods, returns a community partition element that assigns each node to a module/community. In this study a number of algorithms for community detection were evaluated (infomap, Girvan and Newman) (Yang et al., 2016); the one that performed the best on the Çatalhöyük dataset both in terms of modularity score and partitions obtained was the Louvain community detection method. Louvain $Q_{\text{max}}$ was calculated using R (igraph package) while Visone (Brandes and Wagner, 2004) was used to work on the partitions.

**MODULARITY VALIDATION—NETWORK PERMUTATIONS AND ASSORTATIVITY**

For the purpose of this paper, a 2-fold strategy addressing both the $Q_{\text{max}}$ value and the attribution of buildings to modules, has been used to investigate/validate the results obtained applying the Louvain algorithm to the three macro-phase networks at Çatalhöyük.

First, the $Q_{\text{max}}$ values of modularity of the early, middle and late networks are compared against a null model, specifically a “replica” dataset created through a process of edge permutations (randomization) that kept some of key characteristics of the networks constant (size and overall degree distribution) and randomizes the measurements of interest (Croft et al., 2011; Farine, 2017). As previously outlined, according to modularity methods, community structure in an observed network is defined against the modular arrangement of equivalent random networks (RN). The $Q_{\text{max}}$ of the empirical network is interpreted as revealing a significant network community structure if it is substantially larger than the RNs $Q_{\text{max}}$ (Fortunato and Barthelemy, 2007; Fortunato, 2010). In this regard, it should be stressed that large values of modularity don’t always correlate to a significant “community structure” and that networks that are characterized by a clear modular structure display very low values of modularity (Fortunato, 2010, p. 39). Community structure can be present even with very low $Q_{\text{max}}$, though a high value would be preferable (Newman, 2006).

For this study, RNs were created using a process of local edge reshuffling namely randomly rearranging the observed interactions between pair of nodes. In order to create a set of comparable RNs, the permutation process was constrained for both size and degree distribution that are kept the same as the observed network. Permutations were performed in R using the package “tnet” through a link reshuffling procedure (Figure 3; Lusseau et al., 2008; Opsahl, 2009; Opsahl and Panzarasa, 2009; Farine, 2017; Radivojević and Grujić, 2018). The edge permutation process was repeated 1,000 times for each macro-phase network; for each permuted network, $Q_{\text{max}}$ was calculated using the Louvain method (R–igraph package) (Csardi and Nepusz, 2006); frequencies of RNs $Q_{\text{max}}$ were then plotted and compared to the $Q_{\text{max}}$ of the observed networks (Figure 4). The obtained values were very low and not significantly larger than those of the RNs for all three networks. These results come as no surprise given the highly interconnected nature of the networks studied and their expectedly weak community structure. However, a qualitative assessment of the partitions obtained performing modularity optimization leads to the conclusion that even if extremely low, modularity structure is present, and it should be further considered. Chiefly, the strong spatial character of the components isolated by the Louvain method supports a further investigation of the network partitions.

In order to investigate further the modularity results, specifically the assignment of buildings to modules and to verify whether these differ from random, a new set of random networks (100) were generated using a similar process of network edge re-sampling (R–tnet package). The probability that couple of buildings are assigned to the same communities in RNs and in the observed networks is, then, estimated using the assortativity measure on weighted networks (R– assortnet package) (Newman, 2003b; Farine, 2014; Shizuka et al., 2014).

Assortativity has been used to measure the robustness of community structure and quality of sampling strategy in animal societies (Farine, 2014; Shizuka et al., 2014; Shizuka and Farine, 2016). For the purpose of this research assortativity gave us a measure of the similarity of the dyadic associations of buildings to partitions in the empirical network to the one obtained if the network is randomized. Assortativity, in its weighted version, is a value that ranges from 0 to 1, with 1 being the perfect coincidence of buildings to partitions in the RNs and in the observed network and 0 when the attributions are fully divergent (Farine, 2014; Shizuka et al., 2014). In order to obtain assortativity values for the created networks (early and middle) a new matrix was created where 1 is assigned every time dyads of buildings are found in the same community in the 100 replica networks. The membership to modules is then compared to the one of the empirical networks and assortativity calculated.

**RESULTS**

**Early Levels (7100–6700 BC)**

The early levels network is formed by eight buildings (Table 3; Figure 5) confined in a restricted geographical area in the central part of the South Area (Figure 1). The South Area is the part of the East Mound that was originally excavated by James Mellaart during the 1960s (Mellaart, 1963, 1964, 1966, 1967) and that was reopened by the ÇRAP at the beginning of the 1990s. The early network is, therefore, mainly formed of buildings that have been partially excavated by Mellaart and more recently completed by the ÇRAP (B.18, B.43, B.23, and B.2). Additionally, almost all of them, but B.2, are part of deep sequences of overlapping reconstructed buildings whose excavation has been, for the most part, carried out in the 1960s (B.17, B.6, B.24, S.10; B.18, B.16, B.7, B.20, S.8; B.23, B.22, B.21, B.8, S.1) (Farid, 2007a,b,c,d; Hodder, 2007). Of these long sequences the buildings included in the early network have been attributed to three levels (South...
J, South K, and South L) (Table 3) (Farid, 2014a; Hodder and Farid, 2014). They are B.18 and B.23 that are contemporary abutting and communicating buildings that lay directly above the early penning area (enclosed area where very likely animals were kept) revealed through the 1990s deep sounding (Sp.181) (Cessford, 2007), and the two overlapping buildings B.17, B.6. B.17, is similarly to B.18 laying to a possible penning space (Sp.620) (Taylor, 2017). Furthermore, part of the early network is the B.43, B.160, and B.161 sequence fully excavated by the ÇRP in recent years (Farid, 2014b; Taylor, 2017) and B.2 that differently from the other buildings, was constructed over a midden area and never rebuilt (Farid, 2007b).

**Modularity**

The Louvain method of modularity maximization performed in Visone on the weighted early network yielded results on two levels. The first level, that corresponds to the global maximum, is constituted of two clusters (Figure 5B) and the second one which represents the intermediate step in the process of community decomposition, is formed by three modules (Blondel et al., 2008; Figure 5C). The observed early network obtained a very low modularity score ($Q_{\text{max}} = 0.073$) (Figure 4A) that, given the highly interconnected nature of the Çatalhöyük early data is not an unexpected result. The very low $Q_{\text{Emax}}$ speaks to the overall high degree of connectivity of the early buildings and to
the weak pattern of community association and structure. The algorithmically extracted communities are very interconnected between themselves. The very low modularity value is matched by an assortativity value of 0, which means that if we create random networks equivalent to our observed one, we perform modularity maximization on them and compare the probability that couple of buildings belong to the same partition in the RN and in the empirical one, we almost never get the same result value.
Therefore, the observed co-presence of buildings in communities differ substantially from the attribution to partitions in RN.

The first step of the analysis returned two modules (Figure 5B–yellow and pink module) that divide the network into two blocks of buildings B.161, B160, B.6, and B.17 (Figure 5B–yellow module) and buildings B.43, B.18, B23, and B.2 (Figure 5–pink module). The second layer, which is the intermediate step in the modularity maximization appears to provide a good representation of the data and it is made of three modules: the pink module is the same as the previous maximum modularity level, while the yellow one is now divided into two communities (blue and gray modules) formed by couples of overlapping buildings (B.6–B.17 and B.160–B.161). When observing the distribution of material assemblages in the early 2-mode network and in the materials network (Figures 6A–C) it can be observed that the pink module is characterized by the
repeating of specific practice: the location of ovens in the south west corner of buildings together with the presence of underfloor obsidian hoards, the location of burials in the oven area and the quite widespread abandonment practice of putting a projectile point in the pits obtained by the removal of the building posts.

Similarly, the yellow community is defined by recurring practices: the location of a protruding oven in the northeast area of the building, the presence of unusual burial practices, such as including a wooden plank either above or below the body, or the use of pigment, or the presence on the body of a vast amount of small digested rodent bones that were probably introduced in the burials as carnivore scat (Jenkins, 2012). The intermediate modularity layer provides the opportunity to further investigate the yellow module, which is divided at this stage of the analysis into pairs of overlapping buildings: B.161 and B.160, and B.17 and B.6, both share the same footprints and are part of a deep sequence of rebuilds.

Geographical patterns are difficult to observe in this network, nevertheless, it should be noted that in building S.VIII.31 (Figure 5C), the “red shrine” excavated by Mellaart in the 1960s (Mellaart, 1966. p. 180), and abutting to the right the stack of buildings B.43, B.160, and B.161 and possibly contemporary with B.43, shows some of the material associations that define the yellow module that are very rare on site (wooden plank, microfauna, or red pigment in burials). These sets of material practices seem then to define a group of buildings that cluster in space. Regarding the attribution of B.43 to the pink community, it should be said that this building, partially excavated by Mellaart...
as S.VII.27 and left exposed to erosion for a long before being completely excavated by the ÇRP, is a difficult building that weakly connects to the other ones.

**Middle Levels (6700–6500 BC)**

The middle levels at Çatalhöyük are a group of quite homogeneous levels that represent what could be called “classical” Çatalhöyük with its characteristic explosion of symbology within houses. This chronological slice predates the 6500 B.C. period that marks one of the most clearly manifested and widely observed social change at the site (Hodder, 2014a,b). The middle network is made of 18 nodes (Table 3; Figure 7); it is the biggest of the networks under study in terms of both number of nodes and amount of material. Whereas, the early network was spatially constrained to the South Area, the middle one spreads over the South and North Areas and chronologically spans through five levels (Figure 7C). The middle network is almost completely composed of buildings that have been excavated by the ÇRP, seven of which have been fully excavated. Despite its extent, it is the denser and the most uniform of the networks generated and many of the richest in burials and most elaborated buildings that have been studied by the ÇRP are included within this network (e.g., B.1, B.49, B.52, or B.77). The middle network expands mainly horizontally and differently from the early network, few are the sequences of overlapping buildings present, the only exceptions being B.1 and B.5 and B.77 that was, even if much smaller in size, constructed over the earlier B.132. The lack of depth of the network is due to the excavation strategy that was designed to investigate the widest part possible of the site.

**Modularity**

Within the middle network the Louvain algorithm detected one community level made of three distinctive modules (purple, green, and orange modules–Figure 7B). The overall modularity maximum is very low ($Q_{\text{max}} = 0.066$), significantly lower than expected by chance (Figure 4B). The middle network is a dense and highly interconnected network and, similarly to the early network, the very low value of $Q_{\text{max}}$ is not unexpected. The low modularity value is matched by a very low assortativity...
value (0.024). Similarly to the early network the observed co-presence of buildings in communities differ substantially from the RNs partitions.

The green module (Figures 7, 8) is the biggest of the three modules and includes 44% of all the nodes in the network. It is geographically centered in the North Area (B.1, B.3, B.5, B.49, B.59, B.132, and B.114) but one of its nodes (B.50) is situated in the South Area. B.5 is constructed over the footprint of B.1, this is the only overlapping buildings belonging to the same module of the middle network. If we analyze the materials that link these

---

FIGURE 8 | Middle network green module (A), 2-mode network (B – size of nodes according to degree centrality) and object network (C) (The codes used are in the Supplementary Material).
buildings together (Figures 8B,C) we notice that the buildings of the green group share groups of practices: a preference for the location of the highest platform and the burial platform in the northwestern corner of the building, the presence of sub-floor obsidian hoards and of headless bodies in burials. The green module is additionally characterized by a high diversity of chert sources in buildings. B.50 in the South Area connects to the core of the module in the North Area through the shared presence of bird bones in burials and crane bones in the building together with the occurrence of sub-floor obsidian hoards.

The orange module is also centered in the North Area (Figures 7, 9) and is composed of four buildings (B.77, B.131, B.52, and B.96) one of which (B.96) situated in the South Area. If we look at the material assemblages (Figure 9B,C) that keep buildings together in this community, we observe a preference for situating the highest burial platform in the northeast part of the building, all three North Area buildings are burnt and they share a preference for using peas as a legume. B.96 links to the North Area through similar mural painting motifs and burial practices such as the presence of isolated crania in

FIGURE 9 | Middle network orange module (A), 2-mode network (B size of nodes according to degree centrality) and object network (C) (The codes used are in the Supplementary Material).
building and a secondary deposition type that remain quite rare on site.

The third module (purple module) (33%) (Figures 8, 10) is, differently from the other two, centered in the South Area (B.80, B.76, B.79, B.89) only B.102 is located in the northwest corner of the North Area (B.102 is however a problematic building and the attribution to this module very dubious). The buildings assigned to the purple module stretch over three levels: South O/North G and South N. This South Area module is held together by an array of materials, above all specific architectural features and the burnt status at abandonment (Figures 10B,C).

The analysis outlined above highlights several points: first, the middle network is very poorly structured and buildings within it are extensively interconnected and homogeneous in their material assemblages, however the algorithm was able to extract communities that are clearly defined spatially; the three Louvain modules are formed by clusters of buildings adjacent in space. Even so, communities show connectivity that crosscut the local community too. Both B.50 and B.96 connect to building in the North Area through a set of practices that in the case of B.50 are obsidian hoards, bird bones, diversity in chert source and specific chert manufacturing technique and in the case of B.96 regard type of mural paintings and specific burial practices like the presence of isolated heads in burials and secondary depositions (Figures 8, 9).

In terms of connectivity the orange and the green modules (North Area) are the most interconnected while the purple module is the most isolated and homogeneous of the three. This may suggest a difference between the North and the South Areas regarding sets of affiliations. The existence of separate communities inhabiting the two northern and southern prominences of the East Mound has been suggested previously on the basis of skeletal traits (Pilloud and Larsen, 2011).

Late Levels (6500–6300 BC)
The late network is the most problematic of the networks generated in this study. It is the smallest of the three (7 nodes) (Table 3; Figures 11A, 12A–C) and it is still in the process of being integrated. All seven buildings that make up the late network have been excavated by the CRP and the three overlapping ones (B.44, B.56, and B.65) have been fully excavated. Chronologically it develops through five levels (Table 3; Figure 11C) and geographically it stretches to the North and South Areas (Figure 11C). The late chronological grouping follows the changes that mark the 6500 cal. BC. The core of the network is formed by a highly interconnected stack of building (B.44, B.56, and B.65) that were constructed one on top of the other following a similar architectural footprint. This sequence of buildings represents the perfect example of building continuity at Çatalhöyük. B.75 is the earliest of the buildings in the network and, like B.60, is heavily truncated.

Modularity
The community detection algorithm returned one layer formed of three communities (Figures 11B,C). As it will be discussed later the late network modules are not made of clusters of proximate buildings but, excluding the higher interconnected light green module (Figure 11B), the other communities are formed by two buildings each, one located in the South and one in the North area (blue module—B.75, B.58 and pink module—B.42 and B.60). The modularity score (Q_max = 0.039) is extremely low as expected (assortativity has not been calculated given the still transitional nature of this network). If we look at the materials that keep the network together, we realize that only the light green module shares a wide number of practices, while the pink module is very weakly interconnected and B.42 and B.60 share only the presence of secondary burials. B.42 is the building where the only plastered skull discovered at Çatalhöyük has been found (Sadaranangi, 2013). Buildings in the blue module (B.58 and B.75) similarly share only the presence of the evidence of beads manufacture. Both the blue and the pink networks interact more with the main light green module than between themselves.

UNRAVELING THE KNOT: DISCUSSION AND CONCLUSIONS

Within the field of urban studies, Neolithic megasites have been view primarily as anomalies. While they were vast and densely populated settlements, they lack most of the characteristics that have been linked to the classic definition of “urbanity” as it appears in the south of Mesopotamia during the 4th millennium. The density and vastness of their occupation, the fact that these sites were able to sustain highly intense social interactions for centuries through fully integrated social, economic, material and “political” practices, still raises important questions and challenges classic linear models of urbany.

What were the principles of social integration of these large fundamentally egalitarian sites, and what can they tell us about the process of early urbany? While the developments of the Neolithic period were not standardized but instead represent a long-lasting, polycentric and multifaceted process (Gebel, 2002, 2004; Asouti, 2007; Asouti and Fuller, 2013; Finlayson and Makarewicz, 2017), it is nevertheless the case that the Çatalhöyük dataset provides the opportunity to investigate such questions and offers a glimpse of the mechanisms that promoted and allowed the co-residency of a large amount of people in vast early agricultural settlements.

Based on the results of this research, several conclusions can be made. First, the community detection analysis has highlighted a modularity partition essentially based on spatial proximity. Within the highly interconnected system of networks at Çatalhöyük, the modularity maximization algorithm was able to isolate groups of buildings that were related or were interacting more intensely among themselves than with buildings belonging to other sub-communities; these groups (modules/communities) are made of spatially adjacent buildings. Neighboring buildings are embedded in dense sets of material connections that differentiate them from other clusters of adjacent buildings. It should be remembered that, for the purpose of this study, intensity of interaction is measured in terms of similarities of material culture. Therefore, the modularity decomposition analysis highlights that buildings that look alike in terms of
FIGURE 10 | Middle network purple module (A), 2-mode network (B size of nodes according to degree centrality) and object network (C) (The codes used are in the Supplementary Material).
material assemblages and therefore are more intensely connected, are close in space. This is mainly visible in the middle layers where the geographical extent of the network gives us the opportunity of observing it, but it is likewise evident in the early network where communities, nevertheless, are spatially defined and are formed by buildings that are geographically adjacent.

As such, the location of buildings within the site appears to be an important organizing principle at Çatalhöyük. Throughout the duration of the settlement, spatial proximity was very likely to have been promoted actively within the community as a means of organizing social relations and constructing the built environment. A strong commitment to place as a structuring principle is additionally evident in the intense material connections of stacks of overlapping buildings that persists through time. Buildings repeatedly constructed on the footprint of earlier ones are, for the most part, assigned to the same community (only B.43 represents an exception).

Spatial proximity seems to be less of an important settlement-shaping principle in the late levels, although caution is required in the interpretation of the results for this phase given the small sample size. The communities isolated by the algorithm within the late network underscore the strong continuity of connectivity through the three overlapping buildings B.44, B.56, and B.65 and a much sparser geographical communities of buildings. This change in the spatially clustered nature of the algorithmically detected communities might be the result of the less dense built environment that characterizes the end of the East Mound occupation sequence (Hodder, 2013, 2014b; Marciniak et al., 2015). Additionally, the late network is the most fragmented and the least connected of the three constructed networks and when it is binarized most of the nodes get disconnected and almost only the sequence of overlapping buildings (B.44, B.56, and B.65) maintains high connectivity. It should be, however, said that some of the buildings of the TPC, TP, and GDN Areas (Figure 1) that comprise a large part of the excavated late and final Çatalhöyük buildings and that, at this stage of the analysis, couldn’t be added, show some repetition of practices that follow a clear geographical pattern (e.g., Baranski et al., 2015; Baranski, 2016). Among these practices it is worthwhile mentioning the...
habit of inserting an entire pot in the floor in the vicinity of the oven, a practice that is repeated in neighboring buildings B.44, B. 65, B.75, B.42, B.150, B.142, and B.81 and that has not been recorded anywhere else in the site (Yalman et al., 2013; Hodder and Farid, 2014).

Similarities between local groups of buildings have been observed frequently at the site (e.g., Tung, 2013; Yalman et al., 2013; Bogaard et al., 2017) and Hodder (2013, 2014b) lists a vast number of similarities and shared features between nearby buildings; he also suggests that similar groups of buildings might have shared burials location, perhaps within history houses (Hodder, 2014a). Furthermore, neighboring houses show a strong interconnectivity that Hodder (2013, 2014b) refers to possible cooperative practices like herding of animals or hunting of big animals. The same observation is made by Bogaard et al. (2017) when they investigated patterns of legume consumption between adjacent burnt building in the North Area (B.131, B.77, and B.52); in this study they suggest the important role cooperation played in early agricultural societies.

As mentioned previously, the idea of spatially clustering buildings forming neighborhoods as a manner of organizing social relations within the Neolithic dense agglomeration of central Anatolia and beyond, has be suggested before (Düring and Marciniak, 2006; Düring, 2007b, 2013; Hodder and Pels, 2010; Rollefson and Kafafi, 2013). Spatial clustering of structures has been observed at other megaliths and not only at Çatalhöyük. For instance, at the 9th millennium Central Anatolian (Cappadocia) site of Aşıklı Höyük buildings were organized in clearly defined neighborhoods divided by small alleys (Özbaşaran, 2011, 2012; Özbaşaran and Duru, 2015). Aşıklı Höyük predates Çatalhöyük and in many ways anticipates same of the features that would appear later in the Konya Plain such as the clustered nature of habitations and the role of continuity in house construction reflecting a strong commitment to place (Düring, 2005, 2011; Özbaşaran, 2011, 2012).

At 'Ain Ghazal, the best excavated Jordanian megalith, Rollefson (2015) observes that agglomerating houses that appear being occupied by single nuclear families were economically independent, although they probably shared resources within spatially clustered social groups, likely kin-based. Moreover, patterns of habitation based on strong spatial clustering and modular spatial segregation have been observed at other 7th millennium sites like Tell Sabi Abyad (Bernbeck, 2008,

**FIGURE 12** Late network of buildings (A), 2-mode early network (B), and early network of objects (C) (The codes used are in the Supplementary Material).
While modularity at Çatalhöyük is very weak, archaeologists at Tell Sabi Abyad in Syria have discerned a settlement organization based on a sharp spatial segmentation of built structures. It has been suggested that the community at this site and at other large Halaf sites (e.g., Domuztepe, Kazane Hoyuk, or earlier Syrian sites like Tell Mounbatah and El Kerkh) was formed by partially autonomous and kin-based clusters of buildings functioning at the same time but dispersed in the landscape (Akkermans, 2013, 2014). Additionally, it has been possible to observed that within the above-mentioned communities, occupation continually changed and shifted horizontally within the perimeter of the site, creating a vast area of archaeological remains that was however never fully and densely occupied. These fragmented settlements have been able to maintain a strong sense of a united community through their occupation (Akkermans, 2013). This type of habitation has been compared to the later levels at Çatalhöyük in which the settlement becomes less clustered and more dispersed than in earlier periods (Hodder, 2014a). It should be noted that a clear understanding of building contemporaneity and rates of change at Çatalhöyük and other megasites is lacking and that, despite the extensive excavation conducted by the ÇRP, only a fraction of the East Mound stratigraphy has been uncovered (see Akkermans, 2013; Hodder and Farid, 2014; Bernardini and Schachner, 2018).

Some of the adjacent buildings that cluster in the same algorithmically defined community are not contemporary; this is the case, for instance, of B.131 (North Gc) and B.77 (North Gb) in the middle network, or B.132 (North F) and B.1 or B.3 (North G). The groups of buildings detected by the algorithm could point to something different from the presence of tightly clustered neighborhoods on site and on the contrary highlight a more sparse and flexible habitation pattern that was however deeply committed to place and to the repetition of same practices in the same places or the residence of affiliated groups or corporate bodies in the same part of the site. This is evident in the reconstructions of houses in stacks in the same place that represent the persistent search for continuity in social roles and relations and the evidence of history making (Hodder and Pels, 2010; Hodder, 2018; Matthews, 2018).

It should also be said that the groups isolated by the Louvain algorithm could speak to many types of relations between buildings. These groups could be formed by buildings that share a common ancestral affiliation or even a functional affiliation; whichever was the connection between them it had a spatial nature. Geographical modularity, however, is just part of the story; what is really striking at Çatalhöyük is the level of homogeneity and overall intensity of connectivity throughout the entire site. It should be remembered that all three networks buildings are hugely interconnected and the values of community maximization (Qmax) are extremely low and modularity is very poor. There is a marked sense of an overall site-wide shared community identity which appears to be maintained throughout the entire occupation. The middle network is almost impossible to disentangle given the intensity of connectivity. All three networks display affiliation of buildings through shared practices that cross-cut geographical location and Louvain community assignment and bring different parts and groups together. Some of these “linking” practices seem to be the related with mural painting motives (e.g., B.96-orange module), burial practices are both very restricted (e.g., B.6, B.160, B.162–wooden plank in burials or scat of carnivores in burials) and widespread like the presence of isolated heads (e.g., B.96). These types of site-wide connections seem to fit with the proposed idea of the existence of affiliations between buildings as religious sodalities that were bridging different parts of the mound (Mills, 2014). Undoubtedly, we are seeing different forms of affiliation and corporate identities that must have been multifarious and dynamic and that were interacting at different temporal scales (Cohen, 2000; Benz, 2017; Finlayson and Makarewicz, 2017).

In terms of temporal change, it should be stressed that the early network is the easiest to disentangle; furthermore, marked idiosyncrasies are observable in this network when the materials/practices that connect buildings are analyzed. In contrast, the middle network is more homogeneous, and it is extremely difficult to isolate practices and objects that define specific communities. This variability in the early occupation of the site fits well with the a process of community formation and with the idea of the “community born to keep together multiple identities” and possibly different groups coalescing at Çatalhöyük (Finlayson and Makarewicz, 2017). These different identities appear to be homogenized over time, leading to the high levels of interconnectivity observed during the middle network. Spatial proximity might have been sought in order to facilitate a strong sense of community and social cohesion. Through the dense and complex archaeological datasets Çatalhöyük, the network and modularity analyses conducted in this study have provided the opportunity to “discern… the tracery of a pattern so subtle it could escape the termites’ gnawing.”

**AUTHOR CONTRIBUTIONS**

The manuscript has been entirely conceived and designed by CM. CM acquired, analyzed, and interpreted the data. CM, additionally, drafted and wrote the final version of the manuscript and CM is, therefore, accountable for all aspects of the work.

**ACKNOWLEDGMENTS**

I am thankful to Ian Hodder and to all the Çatalhöyük Research Project team members. I am additionally grateful to Francesca Fulminante for inviting me to be part of this Special Issue and for funding the publication of this work. This paper was published thanks to an OPEN-AIRE fellowship granted to Francesca Fulminante for the Marie Slodowska Curie Project 628818 Past-People-Net.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdhig.2019.00008/full#supplementary-material
REFERENCES


Mazzucato

Networks at Çatalhöyük: a Community Detection Approach


Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Mazzucato. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Frontiers in Digital Humanities | www.frontiersin.org 34 May 2019 | Volume 6 | Article 8
The Origins of Trypillia Megasites

John Chapman*, Bisserka Gaydarska and Marco Nebbia

Department of Archaeology, Durham University, Durham, United Kingdom

The Trypillia megasites of Ukraine are the largest known settlements in 4th millennium BC Europe and possibly the world. With the largest reaching 320 ha in size, megasites pose a serious question about the origins of such massive agglomerations. Most current solutions assume maximum occupation, with all houses occupied at the same time, and target defence against other agglomerations as the cause of their formation. However, recent alternative views of megasites posit smaller long-term occupations or seasonal assembly places, creating a settlement rather than military perspective on origins. Shukurov et al. (2015)’s model of Trypillia arable land-use demonstrates that subsistence stresses begin when site size exceeded 35 ha. Over half of the sites dated to the Trypillia BI stage—the stage before the first megasites—were larger than 35 ha, suggesting that some form of buffering involving exchange of goods for food was in operation. There were two settlement responses to buffering: clustering of sites with enhanced inter-site exchange networks and the creation of megasites. The trend to increased site clustering can be seen from Phase BI to CI, coeval with the emergence of megasites. We can therefore re-focus the issue of origins on why create megasites in site clusters. In this article, we discuss the two strategies in terms of informal network analysis and suggest reasons why, in some cases, megasites developed in certain site clusters. Finally, we consider the question of whether Trypillia megasites can be considered as “cities.”

Keywords: Ukraine, Trypillia, Chalcolithic, megasites, settlement structure, assembly places

INTRODUCTION TO CUCUTENI—TRYPILLIA (CT) ARCHAEOLOGY

It seems like a counterfactual proposition that any collection of papers addressing global prehistoric and historic urbanism would be well-advised to heed the forest steppe zone North of the Black Sea in the fifth and fourth millennia BC. For it is in these times in the territory of modern Ukraine and Moldova that you would find examples of the earliest urbanism in the world. In this article, we outline the cultural and social context of what are known as “Trypillia megasites” and discuss the contrasting explanations for their origins.

The Lithuanian prehistorian Marija Gimbutas (1974) coined a phrase for this part of Europe known variously as “Central and Eastern Europe,” “South-East Europe,” and the “Balkans.” Her preferred term was “Old Europe”—that part of Europe with the oldest farming communities and with the closest links to even earlier agro-pastoral groups in the Near East and Anatolia (Figure 1). Gimbutas’ most positive connotation of Old Europe was of a zone connected culturally by shared rich material culture, common ritual beliefs, and a network of matriarchal, matrifocal societies (Gimbutas, 1982). Although “Old Europe” was ideologically created in opposition to the patriarchal Bronze Age (Chapman, 1998), the term is a vivid shorthand for an assemblage of societies which were indeed materially very different from those in Austria, Poland, and points North and West.
One of the leading constituents of “Old Europe” was the Cucuteni—Trypillia group (or CT), recognised as being the “last great Chalcolithic society of Europe” (Monah and Monah, 1997).

One of the most striking characteristics of the CT group was its immense size and chronological depth. The sites of this group covered 225,000–250,000 km², stretching from the Eastern Carpathians in the West to the Dnieper valley in the East, avoiding the North Pontic steppe zone and the East European temperate forest zone to remain within the forest-steppe parkland. Although AMS dating remains patchy, the best estimates for its duration is from 5000 to 2800 BC—how much longer than two millennia remains unclear (Mantu, 1998; Rassamakin, 2012) (Figure 2). No other group in Old Europe reveals such a long tradition, based upon three aspects of material culture—pottery, figurines, and houses. The immense size and the material tradition lasting 65–70 human generations are related insofar as the adoption and millennial continuation of the same material forms in such basic elements of prehistoric lifeways indicates a strong social network that would have attracted the support of communities on the margins, providing a mechanism for continuous spatial growth. We propose that it was the depth and strength of this network that provided the basis for the growth of highly nucleated communities in part of the CT network.

An important result of the spread of CT pottery over such a vast area was the introduction of mixed farming into large parts of the forest-steppe previously settled by hunter-gatherers who made pottery but consumed little domesticated foodstuff (Kotova, 2003). Agro-pastoral communities had been established as far East as the Dniester valley by the 6th millennium BC and, although LBK pottery has recently been found on sites near Odessa (Kiosak, 2017), further discoveries of Trypillia pottery East of the Dniester are assumed to be evidence for the spread of the farming way of life, although whether by movement of people or by assimilation of local hunter-gatherer groups remains unclear. The notion of Trypillia communities as “first farmers” is rarely considered in these debates (but see Müller, 2016b p. 14).

Two related characteristics were shared by both of the CT groups (Cucuteni in Romania and Moldova; Trypillia in Ukraine): the dominance of the domestic, or settlement, domain over the mortuary domain, and the dominance of ceramic over all other forms (metal, stone) of finely made goods. The vast proportion of CT sites are settlements, with no cemeteries known until the very latest phase of the Trypillia group (Phase CII), occasional examples of cave deposition (e.g., the Verteba Cave: Kadrow and Pokutta, 2016) and very few instances of in-mural burial (Bem, 2007). The absence of funerary contexts in which to deposit prestige metal or polished stone items may be one reason for the rarity of metal objects and finely crafted stonework in the CT group. Another reason is what Taylor (1999) has termed “lateral cycling”—the melting and re-shaping of copper into “new” objects. Taylor also argues that ornament hoards constituted a strategy for the defence of valuable copper, as in Karbuna (Dergachev, 1998) and Horodnitsa hoard II (Chernykh, 1992, p. 41). Early CT metalwork was small-scale, regionally specific as to type and rare, often showing signs of repairs (Greeves, 1975; Chernykh, 1992; Ryndina, 1998). Production of larger-scale copper items occurred only from the Middle Phase (BI/II) onwards. By contrast, CT groups produced large quantities of fine pottery which manifested its own special intrinsic value. Painted pottery comprised up to 50% of some Cucuteni Phase A assemblages (e.g., Drăgușeni: Marinescu-Bîlcu, 2000, p. 110).

The third characteristic of the CT group was, in fact, limited to the Trypillia group and concentrated in the Southern Bug—Dnieper interfluve—the growth of the so-called megasites (Figure 1). Megasites were exceptionally large sites of more than 100 ha, with specific planning features such as concentric circuits of houses and a large, open inner space (Videiko, 2013). From the late 5th millennium BC onwards, a divergence trajectory in settlement size and nucleation separated Cucuteni from Trypillia. In the Cucuteni A phase, settlement numbers increased as size fell to a mean of 1 ha, with a resultant dispersion of settlement across the landscape. A good example consists of the Cucuteni settlements in Bacău County, North-East Romania, in which small sites spread from the main valleys into third- and even fourth-order stream catchments (Popovici, 2000, Figure 2). The opposite development occurred in the Trypillia A phase, with 1-ha sites still found but occasional nucleated sites such as Mogyla III reaching 10 ha in size (Videiko, 2007, Table 1). Increased nucleation is seen against a background of the continuing dominance of small sites in the Phase A-BI transition (Steapanivka: 15 ha), Phase BI (Chyzivivka: 20 ha) and the BI-BII transition, with several sites larger than 100 ha (e.g., Vesely Kut, Kharkivka) and even sites of up to 200 ha claimed (e.g., the eponymous site of Trypillia). The strong trend toward settlement dispersion in the Cucuteni area is a very good reason for the absence of mega-sites in Moldavia—but why did the opposite occur in the Trypillia zone?

**TRYPILLIA MEGASITE INVESTIGATIONS (TABLE 1)**

The investigations of Trypillia megasites forms part of the later development of Trypillia research, from the 1960s onwards. Following Kuhn (1970) model of revolutions in scientific knowledge, we have divided megasite investigations into two revolutions, each followed by periods of “normal” archaeological research (Table 1).

The “second methodological revolution” (Chapman et al., 2014a) led to a new generation of much more accurate geophysical plans which revealed a wide range of new plan features and combinations of features at megasites such as Nebelivka (Chapman et al., 2014a), Majdanetske, Taljanki, and Dobrovody (Rassmann et al., 2016a). The Nebelivka project focused on the integration of a wide range of data lines to

---

1“Tripolye” in the Russian literature.

2For more detailed accounts of the development of megasite archaeology, see Gaydarska (2019a), section 1.1.
provide a challenge to the traditional account of megasites as permanent settlements with thousands of people (the “maximalist” view: Müller et al., 2016a) The combination of nine different lines of evidence produced a “tipping point” in megasite interpretations (Chapman, 2017), which led to three alternative models of smaller-scale, sometimes seasonal settlement models—the Distributed Governance Model (Gaydarska, 2019b), the Assembly Model (Nebbia et al., 2018) and the Pilgrimage Model
TABLE 1 | The main stages of investigation of Trypillia megasites.

<table>
<thead>
<tr>
<th>Stage of investigation</th>
<th>Key characteristics</th>
<th>Site examples</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery stage (1890s–1900s)</td>
<td>Discovery of Trypillia sites defined by burnt houses; comparison of painted pottery to other European Neolithic painted wares</td>
<td>Trypillia</td>
<td>Krivoika, 1901; von Stern, 1900</td>
</tr>
<tr>
<td>1st period of “normal” excavation</td>
<td>Discovery of hundreds of new Trypillia settlements; excavation of representative samples</td>
<td>Vladimirovka/Volodimyrvka</td>
<td>Passek, 1949</td>
</tr>
<tr>
<td>1st methodological revolution</td>
<td>First aerial images of megasites; ground-truthing of megasites; first geophysical investigations, with targeted excavation of house-shaped anomalies</td>
<td>Taljanki; Majdanetske;</td>
<td>Duckin, 1978; Ellis, 1984</td>
</tr>
<tr>
<td>2nd period of “normal” excavation</td>
<td>Large-scale excavations on two of largest sites; geophysical plans of other sites; refinement of Trypillia ceramic typo-chronology (Ryzhov)</td>
<td>Taljanki; Majdanetske;</td>
<td>Shmagli and Videko, 1990; Kruts, 1990</td>
</tr>
<tr>
<td>2nd methodological revolution (2009–present)</td>
<td>Improved geophysical methods, leading to more accurate plans; discovery of new features (assembly houses, pits, kilns, ditches, paths) and groups of features; use of AMS dating, pollen and phytolith analysis; spatial analysis of megasite plans;</td>
<td>Nebelivka; Taljanki; Majdanetske; Dobrovody; Apolianka</td>
<td>Chapman et al., 2014a; chapters in Müller et al., 2016b; Hale et al. (2017); <a href="https://doi.org/10.5284/1047599">https://doi.org/10.5284/1047599</a> ADS YORK sections: Hale, Millard, Albert Johnston</td>
</tr>
</tbody>
</table>

(Chapman and Gaydarska, 2019). While each of the three models is informed by contrasting decisions about seasonality and building strategies, they share many communalities in the reasons for megasite origins.

TRADITIONAL “MAXIMALIST” ACCOUNTS OF THE ORIGINS OF MEGASITES

The principal source of complexity in the Trypillia group is the unique incorporation of elements of two of Gordon Childe’s “Revolutions” in the same group. While the spread of CT documents the spread of the Neolithic Revolution, the development of Trypillia megasites illuminates aspects of the Urban Revolution. Unlike most other regions in the world, these developments are separated by only one millennium. It will be important to distinguish the effects of the two Revolutions in any discussion of megasite origins.

In a paper entitled “Two studies in defence of migration concept,” Dergachev (2002) documents the spread of the use of CT pottery—read as people—across the forest steppe zone, showing in a series of maps the 5-fold sequence of core settlement zones and expansions into hunter-gatherer lands (Dergachev, 2002, Figure 6.2a–e). Waterbolk (1968) notion of the huge reservoir of Holocene soil fertility available for the LBK first farmers in Central Europe applies just as effectively to the chernozems of the Ukraine—some of the richest soils in Europe (Kubiena, 1953) and surely offering huge land-use potential to Trypillian first farmers. However, the intriguing fact is that Dergachev never once mentions the impact of these migrations on the formation of megasites. Rather, population movement was a response to the widespread availability of free land, which continued into the Late Trypillia phase in significant areas, as well as to military threats (see below).

More recently, Diachenko (2012) has invoked population pressure in the form of a population boom in the BI phase to account for the formation of early megasites. He relies on exactly the same site data as Dergachev (2002; compare Figures 6.2a with 6.2b)—population migration into the Southern Bug—Dnieper Interfluence from the Dniester valley—but with the introduction of site population estimates. Diachenko and Menotti (2012) have used the gravity model to trace “genetic ties” between pairs of sites in the Bug—Dnieper Interfluve through time, based upon Ryzhov’s typo-chronological method (Ryzhov, 2005, 2012). However, Diachenko & Menotti fail to explain why such migrations led to the creation of megasites rather than just village clusters in areas of high arable potential (cf. Diachenko, 2016).

One well-known advantage of settlement nucleation is the protection it affords residents in crises of internal or external aggression and warfare (Chapman, 1988; Müller, 2016a).

Could the positive feedback cycle of increased settlement nucleation—greater threat from larger armed groups—even more nucleated defence have led to the trajectory of increased Trypillia site size discussed above?

Echoing Chernysh (1977) and Gimbutas (1977), Kruts (1989, 1993) argues that the principal threat to Trypillia communities came from the Sredni Stog groups in the steppe zone to the South and East, which is why the greatest concentration of megasites was located near the forest-steppe—steppe border on the Southern side of the distribution. However, to the extent that even 10–20-ha Trypillia sites would have been large enough to deter armed Sredni Stog raiders, there was no military reasons for much larger agglomerations—and certainly not for sites of over 100 ha.

Dergachev (2002) supports the view of a steppe invasion with his finding of a higher ratio of fortified to non-fortified sites, and higher numbers of arrowheads per site, in Phase BI than in Phase BII. He suggests that Phase BI was a “society... literally under siege” (Dergachev, 2002, p. 103), in a “state of war owing to outside threat” from the steppe (Dergachev, 2002, p. 106), contrasting Phase BII as a period of relative peace, with the removal of siege and military threat (Dergachev, 2002, p. 103). While this view can be used to support the appearance of early (BI) megasites, it offers no support for the military explanation for the largest megasites of Phases BII and CI.
By contrast, Videiko (2007, p. 274–5) proposed an internal social conflict for the origins of megasites, describing Trypillia chiefdoms as “in a state of perpetual internecine war” (cf. Dergachev’s view but for a later Phase) because of the expansive nature of Trypillia agriculture, with each site exhausting their local soil potential every 40–70 years and needing to move on to capture more arable land. Even if the maximalist assumption of massive megasite populations was not met, Videiko ignores the large unsettled areas in the Southern Bug—Dnieper Interfluve, even in Phase BII (Figures 3C, D). There is also little evidence for warfare, with two exceptions. At Drutsi I, in Moldova (Ryndina and Engvatova, 1990), lithic distributions showed an archery attack on a small site. More compelling evidence derives from the Verteia Cave, where 11 out of 25 buried crania have clear indications of trauma (Madden et al., 2018). However, none of these crania has been directly dated and the site is far from any megasite, thus jeopardizing any potential link between the phenomena.

It is clear that migrations can provide a method for moving people across the landscape but not a reason for any particular settlement form—say, megasites rather than village clusters. This leaves internally-driven or externally-imposed warfare as the principal traditional explanation for the rise of megasites—not the outcome predicted by Gimbutas (1977) peaceful matriarchal CT society!

Many of the problems with these traditional explanations are tied to basic maximalist assumptions about the megasites themselves. Once the population estimates of tens of thousands of people on a megasite are accepted, large-scale processes are required to conjure up the masses. This usually involves grade-inflation: bigger-than-usual migrations, sustained baby booms or mega-battles³. The fundamental underpinning of these explanations—especially the modelling—is Videiko (2002) claim for the coeval dwelling of as many as 78.4% of houses on a megasite (see also Müller and Videiko, 2016). Once this claim is challenged (Chapman, 2017; Chapman and Gaydarska, 2019), new possibilities open up for the debate on megasite origins. In the first part, we discuss alternative readings of the settlement and subsistence evidence, before turning to tradition and innovation in Trypillia material culture.

ALTERNATIVE EXPLANATIONS I—SETTLEMENT AND SUBSISTENCE

There are two basic issues with discussing Trypillia settlement—a paucity of intensive, systematic fieldwalking programmes and the lack of a critical appraisal of existing settlement data (Nebbia, 2019). Nebbia’s filtering of the settlements listed in the “Encyclopaedia of Trypillia Civilization” (Videiko, 2004) reduced the number of sites with clear location, size, and Phase information from over 2,500 to just under 500 (Nebbia, 2019). Equally, the fieldwalking programme for the Nebelivka Project led to the discovery of two new Trypillia sites in a surveyed area of 15 km². Since the Bug—Dnieper Interfluve—an area of c. 50,000 km²—has received hardly any detailed fieldwalking coverage, the trends discussed here can be little more than preliminary suggestions.

The spatial distribution of sites in the Dniester-Dnieper interfluve suggests different levels of clustering/nucleation from the Forest Neolithic phase onwards (Figure 4), and therefore a consideration of second-order effects of the site distribution will help in clarifying social relations between sites. A Ripley’s K-function was used in order to explore the clustering at different scales across the four phases (Ripley, 1976). In Figure 3, plots representing K-functions are shown for the four point patterns (Forest Neolithic, Phase A, Phase BI, Phase BII). These plots display the expected values of complete spatial randomness (CSR) (K_{obs}(r)) and the observed values (K_{obs}(r)) where r represents distances between points. The K (r) values were estimated for 999 random Monte Carlo-simulated patterns and compared with the values estimated for each dataset (Baddeley et al., 2014). A Ripley’s isotropic correction was adopted in order to reduce the edge effect (Oswar, 1983).

If the K (r) is higher than the top of the Monte Carlo envelope, it means that, at that distance, the points are clustering and the hypothesis of spatial randomness can be rejected. Figure 3 shows the progressive diachronic increase in the scale at which sites are clustering, even at short distances (5–10 km) in Phase BII. For the earlier phase of hunter-gatherer settlement, the hypothesis of complete spatial randomness cannot be rejected as the observed values remain within the simulated envelope. In the Trypillia period, there is a significant increase in spatial interaction at short distances for sites in the Southern Bug—Dnieper interfluve, meaning an underlying process of site clustering. The identification of these clusters was facilitated by a Kernel Density Estimation (KDE) for the four point patterns (Figure 5). Although the K-function suggested that a complete spatial randomness could not be rejected for the Forest Neolithic groups, the plot shows how around 20 km the K (r) values are higher than the simulated envelope, and therefore a minimal spatial interaction is occurring (Figure 5A). For the Trypillia phases, it is clear that increasing numbers of cluster were co-emerging with the mega-sites themselves (Figures 5B–D). Moreover, the KDE plot shows how the clusters themselves show an overall aggregation within the wider area, thus suggesting an even higher degree of interaction at a larger scale between different site clusters.

The basis for a discussion of Trypillia settlement is the trajectory toward nucleation at selected sites from Phase A onwards in the Southern Bug—G. Tikych river system. It is important to note that Phase A site clusters were located along the Southern Bug in areas of traditional hunter-gatherer site groupings (viz., Forest Neolithic sites: Gaskveych, 2019, Figure 5.29: here, Figure 5A), indicating long-term continuity in favourable settlement locations. Phase A settlements were strung along the Southern Bug like beads in groups of up to five sites, including the largest sites—Mogyla III and Stepashnya—both already large sites and in different site clusters. One site in the G. Tikych valley was settled in Phase A. As with the hunter-gatherer groups, the network of smaller streams was generally avoided (Figure 5B).

³Perhaps the extreme size and duration of the CT group provides implicit support for such mega-ideas.
FIGURE 3 | Plots of the K-functions for (from Top to Bottom) Forest Neolithic sites (N = 27), Trypillia A sites (N = 33), Trypillia BI sites (N = 46), Trypillia BII sites (N = 176). Distances r are in metres. The envelope has been generated from Monte Carlo simulation (999 iterations) under CSR.
Settlement in Phase BI showed a combination of continuity and expansion. The same three site clusters were occupied along the Southern Bug but there was a major expansion along the network of small streams (Figure 5C). However, all three large sites in the Phase BI and BI-II transition (Chyzhivka, Vesely Kut, and Kharkivka) were located in the same site cluster in a main tributary—the upper part of the G. Tikych valley. The discovery of traces of casting and production waste alongside earlier production methods indicates extractive metallurgy at the largest early megasite—Vesely Kut (Ryndina, 1998, p. 136–150 & Ris. 66/19).

A major expansion along small stream networks was the defining characteristic of Phase BII, with its many new settlement clusters and growth in megasite size (Figure 5D). Despite the continuation of settlement in one site cluster upstream on the Southern Bug, settlement changes can be seen in the abandonment of the longest-lasting site cluster on the Southern Bug and the opening up of new site clusters both along the upper parts of the G. Tikych and along many small streams. The location of the first megasites adjacent to smaller streams can be dated to this Phase.
What this long-term settlement pattern indicates is the establishment of solo settlements before the emergence of a new site cluster in the succeeding Phase—a well-known pioneer colonising strategy (Anthony, 1990). An important development is the inclusion of sites much larger than the usual in two of the clusters. This dwelling strategy led to a growing number of site clusters in the Southern Bug—G. Tikych system, some of them including early (BI and BI-II transition) megasites. What can account for the emergence of site clusters?

The process of farming groups dwelling in a relatively unfamiliar terrain populated by hunter-gatherer populations in main valley site clusters would have required two contrasting settlement choices—proximity to hunter-gatherers for peaceful interaction and distance from hunter-gatherers for security. One way to achieve both goals was the creation of small site clusters near to the hunter-gatherer locations. The emergence of a single large site in such agro-pastoral clusters would have intensified interaction over several farming clusters as well as being attractive to hunter-gatherers.

Another benefit of site clusters was the buffering opportunities offered by kin-related communities in case of crop failures or poor harvests (Halstead and O’Shea, 1981). The argument is that long-term exchange networks between nearby communities would provide security through additional food exchanged for desirable goods such as fine pottery, high-quality flint, copper, or polished stone axes. However, such buffering may not have been so important in Phase A owing to three factors: (1) the small size of settlements, which (2) put little pressure on local chernozem resources, whose (3) Holocene fertility reserves had scarcely been touched.

It was only with increases in settlement nucleation in Phase BI that the opportunities for buffering may have become significant, when the sharing and exchange of resources without the need for a structured socio-economical organisation to regulate the network would have stimulated looser inter-kin interactions, with less resultant social pressure. Shukurov et al. (2015) have modelled the agro-pastoral potential of Trypillia landscapes, reaching the conclusion that the local soil and forest resources were capable of supporting settlements up to the size of 35 ha. However, site clusters in the same areas may have begun to put pressure on even the legendary fertility of chernozems. Moreover, BI and BI-II settlements were growing to a size well beyond 35 ha—indeed to 100 ha and over. Apart from the solution of using only a part of the houses at such large sites at any one time (see above, p. 3), a more complex intra-cluster practice may have involved the provisioning of the largest sites from smaller settlements in exchange for ritual services and exchange items. The site clusters could thus have opened up a space for inter-site functional differentiation involving ritual leadership and the transfer of food and drink to such centres. It is suggested that this scenario may have kick-started a long-term role of assembly places in Trypillia site clusters, at least...
partly based upon the strong social networks connecting local and more distant settlements (see below, pp. 7–8). However, it is still a long way from Phase BI assembly places to BII megasites such as Nebelivka and CI megasites such as Taljanki and Majdanetske. How did this trajectory take root and progress?

The size of the overall Trypillia group is such that we have to assume the development of inter-site interactions over a much greater distance than in other groups (e.g., the Csőszhalom group: Raczky et al., 2007). A significant change would have been the foundation of an assembly place which attracted people from more than one site cluster. What was the scale of attraction of early megasites?

An additional analysis of second-order effects was conducted on the spatial behaviour of values of site size within the whole Trypillia period, thus including Phase CI and CII data, for a total number of 499 sites with good-quality information. An incremental Global Moran’s I (Moran, 1950) on site size values has been calculated for 30 iterations of five site distributions (one for each Trypillia phase), starting from an initial distance band based on the 2nd nearest neighbour count in order to test the scale of site size clustering. Using the chronological phases as time blocks, the results showed how the onset of clustering during phases BI, BII, and CI at 84 km – 93 BI sites, 112 km – 176 BII sites, and 100 km – 236 CI sites. The scale of ∼100 km becomes meaningful when it is constant for the duration of mega-sites occupation of approximately 1,000 years. A LISA (Local Indicators of Spatial Association) test (Anselin, 1995) supported the hypothesis that mega-sites are outliers and provided further confirmation at 95% confidence⁴ that these are outliers of high values within a 100 km neighbourhood of low values. An interesting result is that megasites from the Southern Bug – Dnieper Interfluve had overlapping catchments of 100 km, which might suggest the competitive nature of megasite interaction in that area.

In fact, the 100-km scale of interaction meant that there was no reason why an assembly place of sufficient reputation could not have attracted participants from another site cluster in Phase BI. In Phase BII, the close proximity of site clusters across the Southern Bug—Dnieper interfluve reflexively created the opportunities for visits between site clusters, with all the attendant social potential for significant growth. But we are still far from the typical megasite planning elements that have defined megasites since their discovery and even further from an account of the cultural foundation of Trypillia social networks. A background narrative for settlement history is a necessary but insufficient story to provide a convincing explanation of megasite origins.

⁴For a full methodological explanation see Nebbia, 2017.
ALTERNATIVE EXPLANATIONS II—TRADITION vs. INNOVATION

The Possibility of a Megasite

Before further discussion of alternative trajectories toward megasites, we should step back and consider one fundamental issue. In his influential study of Imagined Communities concerning the anomaly of modern nationalism, Anderson (1991, p. 4) reminds us that all communities larger than a single village are “imagined communities.” By implication, we suggest that integration of people beyond their normal, face-to-face groups required a vision of how those diverse communities could live together to derive benefits from the new settlement form that were considered greater than the difficulties this linkage may have brought. After all, there is a long tradition, beginning with Childe (1958), of praising the advantages of autarky—living in independent, face-to-face communities—a strategy which has, by and large, limited the scale of settlement nucleation in prehistoric Europe. Nonetheless, the existence of the Trypillia megasites is an obvious negation of small-scale communities; their scale and size engenders an equally sizeable problem of how such communities were imagined in the first place.

For let us be under no illusions: on the Eurasian continent of the 5th—4th millennia BC, the Trypillia megasites were unique in size and scale. There was nothing anywhere else on the planet to compare with the Phase BI megasite of Vesely Kut, covering an area of 150 ha—no analogies from which to derive this extraordinary place. We should never forget the unprecedented nature of Trypillia megasites, which have created immense problems of explanation and understanding but, first of all, problems of imagination. A better understanding of this issue comes from defining what social relations were in place before the imagining and the form of these relations‘ materialisation—whether objects or site plans. In this section, we consider how existing elements known to Trypillia communities were juxtaposed and combined in a process known as “bricolage.” This anthropological term signifies the construction or creation of a work from a diverse range of things that happen to be available. Used by Levi-Strauss (1962) to refer to the process of myth-making, bricolage was extended by Derrida (1970) to refer to any form of discourse. We consider the Trypillia Big Other, inter-regional exchange networks and the development of settlement planning as three critical bricolage-led contributions to the emergence of megasites.

The Trypillia Big Other

The massive size and great temporal depth of the CT group was founded upon a strong social network connecting communities at both the local and the regional level. We have previously discussed the importance of what we term the “Trypillia Big Other” for integrating the vast number of Trypillia settlements and their residents. We think of the Big Other as a suite of beliefs which was materialised in practices involving the three key Trypillia traits—houses, pottery and figurines (Chapman and Gaydarska, 2018a; Gaydarska, 2019a) (Figure 6). The term “Big Other” was developed by Lacan (1988) and elaborated on by Žižek to convey the sense not of an ideology nor a religion but an effective symbolic fiction playing a significant role in everyday life (Žižek, 2007a,b). Kohring has discussed the Big Other in terms of its impact on the Bell Beaker assemblage, acting as “a material/symbolic mediator for a whole network of shared conceptual structuring principles” (Kohring, 2012, p. 331). One of the greatest attractions of the Big Other for us is that it is “something which is sufficiently general and significant to attract the support of most members of society but, at the same time, sufficiently ambiguous to allow the kinds of localized alternative interpretations that avoid constant schismatic behaviour” (Chapman and Gaydarska, 2018a, p. 267). Thus, the Big Other has allowed myriad regional and local variations in house-building, pottery, and figurine production yet, all the while, retaining an overall attachment to Trypillia identity. Bricolage was involved through the selective permutation of different elements of the Big Other to produce local forms, with their attendant practices, best suited to the local community without straying too far from overall principles.

However, major changes occurred at the transition from Phase BI to BII in the ceramic aspect of the Big Other. Although painted pottery was the predominant fine ware in North-East Romania and Moldova in Cucuteni Phase A (Popovici, 2000), it was rare in comparison with incised wares in Phase BII in the Southern Bug—Dniester Interfluve (Palaguta, 2007). The spread of trichrome painted wares, with red motifs outlined in black on a light background, characterised Phase BII in this area, providing a novel medium for household identity and linking settlements in a developed version of the Big Other. The assessment of the importance of this ceramic innovation to megasite origins remains an urgent task.

It is therefore hardly credible to us that megasites could have emerged without the mediating, integrative potential of the Big Other to provide the basis for everyday social practices on all Trypillia settlements, viz. the habitus (Bourdieu, 1977). In the context of megasite origins, shared participation in the Big Other and its quotidian materialisation created pre-existing bonds between communities in different sites living in different site clusters, often quite remote from each other. It was the Big Other that reduced the social difference between communities separated by much physical space, providing common grounds for meeting strangers as well as brothers on assembly places. But the bricolage of the many varied elements of the Big Other also enabled communities to create difference without threatening either inter-site relations or local community identities.

Trypillia Exchange Networks

The second part of the ancestral past which Phase BI and II communities relied upon to create megasites consisted of the pre-existing exchange networks. Most Balkan Neolithic and especially
Chalcolithic communities played important roles in often long-distance exchange networks featuring copper, gold, obsidian, and flint, polished stone of many kinds, marine shells such as *Spondylus* and finished objects such as pottery, ornaments and other prestige goods (Chapman, 2019). There have been many CT sites with the deposition of objects or materials exotic to the CT distribution.

A degree of network continuity is demonstrated by the exchange of the lithics essential to many maintenance activities on any Trypillia site. However, the large fall from thousands of items deposited in Phase A and BI sites to hundreds on Phase BII sites (Kiosak, 2019) was a major and as yet unexplained change. All of the BII and CI megasite lithic assemblages so far analysed have included a sizeable proportion of high-quality flint (often up to 50%) from the Prut—Dniester valleys, indicating exchange over 200–300 km.

The major changes in exchange networks concerned manganese and copper. The black pigment manganese was essential for Trypillia painted vessels in Phases BII—CI—a high-value, low-bulk material with sources in the East Carpathians, the Lower Dnieper valley, and the Crimea. The most recent characterisation studies confirm Ellis (1984) identification of the main sources in the Eastern Carpathians (Buzgar et al., 2013), indicating low-bulk, high-value exchange over 300–500 km. This aspect of Trypillia exchange hardly touched Phase BI sites but was vital for BII settlements. There was also a major re-orientation of copper exchange networks at the start of transitional Phase BI-BII, with sources in Transylvania preferred to the hitherto dominant Bulgarian sources (Ryndina, 1998). The question of high-bulk, long-distance salt transportation from either the Eastern Carpathian sources or the North Pontic limans remains under discussion (Chapman and Gaydarska, 2003; Mircea and Alexianu, 2007).

However, when we turn to prestige goods, there is something of a “white hole” for exotic copper or polished stone items in the Southern Bug—Dnieper interfluve. A very rare *Spondylus* bracelet in Lysaya Gora, in the Lower Dnieper valley, has good stylistic parallels with the West Pontic Chalcolithic cemeteries at Varna and Durankulak (Chapman, 2002) but no such marine shell finds are known from the megasites. Equally, the serpentine bracelet from the pre-Caucasus range deposited at Novi Rușești in Moldova has no parallels in megasite deposition. There has been no analysis yet of the only gold ornament yet found on megasites—the gold spiral at Nebelivka (Chapman et al., 2014c, Figure 17).

To summarise this complex data set (Figure 7), all Trypillia settlements in the Southern Bug—Dnieper Interfluve would have required lithic raw materials for basic tool-making—whether from local quarries or exotic sources in the Prut—Dnieper valleys. Local sources would also have supplied stones for grinders and mortars. While there were widespread local sources

---

7One example of the few copper objects from a megasite is the copper axe from Majdanetske House Zh-2 (Shmaglij and Videiko, 2002, Figures 54/15 & 55/1).
for red, white, and orange pigments, black pigments from Phase BII onwards was an exotic for the Interfluve, probably from the Eastern Carpathians. Transylvanian copper would also have been transported across the Eastern Carpathians. Thus, exchange of exotic flint, copper, and pigment alone would have been predicated upon an inter-regional network connecting dozens if not hundreds of sites—a network which would have been instrumental in the consolidation of the Trypillia Big Other as well as maintaining contacts between neighbouring and distant communities. An inter-regional network for exotic lithics would have been operational in Phase A, with an expansion in Phase BII to transport manganese for pot-painting and Transylvanian copper. The paradox of Trypillia exchange dates to Phases BII and CI—the peak of the megasites—when the expected social differentiation consequent upon the development of such massive sites fails to find materialisation in exotic prestige goods on the megasites themselves. This is all the more surprising when we recall that exotic prestige goods exchange was one of the foundations of the Balkan Climax Copper Age. Is it possible that we have grossly over-estimated the significance of Trypillia exchange? Or does lateral cycling hide the multiple re-working of copper objects—the first such recyclable material in prehistory?

Trypillia Settlement Planning

If the Trypillia Big Other provided the necessary material constancy in a cultural tradition and inter-regional exchange networks maintained links between communities through the transmission of ideas, materials and marriage partners, the evolution of planning on Trypillia settlements provided the spatial context for megasite living. The megasites were not only about size, although this was key to their significance—they were also concerned with spatial order and the provision of structure for such huge settlements. Trypillia megasites were based upon the principle of concentricity—unlike the Balkan tell principle of Singerei, Moldova, show only weak tendencies to house concentricity and no inner radial streets or open inner space (Rassmann et al., 2016b, Figure 6) (here Figure 8). However, a careful re-examination of the plans of Phase A, BI, and BI-BII transition megasites shows that not one single early megasite contained all of the four key planning principles of the developed megasites—rather, they rarely contained more than one element. Instead, many of the early megasites contained house nests and concentric house nests that typified Cucuteni settlements as a “hang-over” from pre-megasite planning (e.g., Trușești and Hăbășești: Popovici, 2010). This crucial finding underlines the variability which one may expect in megasite plans of the BI and BI-II Phases. It also shows that, rather than inheriting the blueprint of a complete megasite plan, planner-builders of BII megasites such as Nebelivka improvised a complete plan with all four planning elements as they built the site (Figure 9). This form of bricolage is typical of cultural creation based upon improvisation rather than faithful copying of a pre-existing design. It was not that the planner-builders of Phase BII megasites had nothing to use in formulating a site plan—rather that decisions taken in the process of creating a site were taken based upon a combination of cultural memory and direct witness. This result emphasises the creative bricolage of the BII megasite planner-builders in forming a fresh, previously unknown megasite plan from elements selected from the ancestral past. The result was the spatial formalisation of an assembly place in terms of the two principal spaces—the outer space for dwelling and the open, inner space for assembly. It is suggested that the formalisation of megasite planning in Phase BII was a vital advance toward megasite development, which allowed the evolution of Phase CI sites of even greater size and complexity. Moreover, this advance also influenced the new formalisation of the layout of smaller settlements. An example shows how the CI site of Apolianka (7 km West of Nebelivka—Figure 10) reproduced on a much smaller scale two of the four key elements of megasite planning: a house circuit defining a central open space.
Gatherings of different group sizes must have taken place well before the emergence of megasites so as to underpin the cultural uniformity of CT. While small and medium-size settlements would have comfortably accommodated a local gathering of settlements within a 30–40 km catchment, intra-regional or inter-regional assemblies of 100 km would have required a much larger space. The accumulated experience of the benefits of such gatherings—a substantial increase of opportunities for social interaction, access to "exotic" goods, scaled-up rituals, feasts, and ceremonies, etc., together with the efforts to "set up" and manage such massive aggregations, may have led to the realization that they need not be always temporary or that organization of such events should always start from scratch. The formalization of the best of both worlds—the space for large gatherings and the everyday habitus—reinforced the accumulation of place-value through the incorporation of two very important social principles in CT lifeways. These novel aspects of settlement planning are part of a new knowledge that developed within the experience.
of the making of megasites and which broadened the shared material practices that constituted the Trypillia Big Other. The changes in a dynamic social milieu which allowed for megasites in the first place also could have led to disputes and breaks in former alliances, stimulating the founding of alternative assembly places, which would have led to competition between emergent megasites, even in the same site cluster.

**DISCUSSION**

The possibility of a Trypillia megasite was not an on / off possibility but a contextually rooted concept always *in statu nascendi*, depending upon the potential of the forms of settlement plan, exchange networks and Big Other known at the time. Far from seeing it as in martial crisis under a state of siege (Dergachev, 2002), we think of Phase BI in the Bug—Dnieper Interfluve as a time of both settlement consolidation in the main valley site clusters of Phase A and settlement expansion into the network of smaller streams which defined plateaux and promontories for dwelling. The emergence of settlements larger than the 35-ha. threshold of local sustainability (Shukurov et al., 2015) was limited to one site per cluster in the main valley site clusters in the G. Tikych valley, with smaller sites in the smaller valleys. These earliest megasites had begun to create concentric house circuits and inner open spaces in their plans, alongside the traditional house nests of Trypillia Phase A and indeed much of Cucuteni settlement planning. Phase BI site plans had by no means coalesced into a settled planning system (*Figure 8*)—a development not seen until BII megasites such as Nebelivka—but were creating dwellings with an unprecedented scale and number of inhabitants.

It is hard to conceive of successful attempts to integrate so many people at megasites without an early version of the Trypillia Big Other—the Phase A version, accepted by most people in most former and existing settlements. Nevertheless, we should not forget the fundamental changes to the Big Other, notably the innovations of painted pottery and figurine styles, that were occurring during Phase BI—at the same time as major changes in settlement form. Both types of objects offered new resources for identity-formation in times of immense change except in one key area—dwelling houses. There is remarkable continuity over the whole CT distribution in house-design (Burdo et al., 2013)$^8$, the context for family living which underpinned the dwelling process of Trypillia settlements. The mutual reinforcement of the Big Other by inter-regional, regional, and local exchanges of stone, pigments and metals strengthened inter-community ties in ways that were particularly important at the local dwelling level. Supplying each site with basic local stones for grinding grain and making cutting and scraping tools tied communities into a landscape routine and a set of social relationships for sharing the stone between houses (Skourtopoulou, 2006). The use of exotic flint from the Prut—Dniester valleys not only linked the people in the Bug—Dnieper Interfluve to their Western roots but provided the means for differential acquisition of high-quality flint. It seems that lithics formed the basis for regular, repeated inter-site exchange, with the movement of finely-crafted stonework and marine shell ornaments a far more occasional practice probably "piggy-backing" on pre-existing lithic, copper, and pigment exchange networks.

We return to those Phase BI settlements which transgressed the 35-ha. threshold of local sustainability. We submit that no-one has yet provided a well-documented case of the coeval use of 80%, or indeed 100%, of a megasite’s houses$^9$. One solution to the problem of sustainability is the acceptance of a small fraction of houses in coeval use—in the case of Vesely Kut, perhaps a quarter of its houses, whereas a third of its houses at the smaller Kharkivka.

Another solution—by no means incompatible with the first—concerns the stimulus of new social relations between the largest and the smaller sites in the site cluster. These relations provided a

---

$^8$This is not to deny that regional differences in house designs were present (Burdo et al., 2013).

$^9$Attempts include Videiko, 1996; Diachenko, 2012; Müller and Videiko, 2016; Müller et al., 2016a. For a detailed critique, see Gaydarska, 2019a Chapter 6.
form of buffering which was hardly necessary in Phase A. The co-emergence of the growing size of a megasite with its reputation as a ritual and exchange centre led to a synergy between locals and other residents in the site cluster. The provision of food, drink and possibly other resources (such as salt or copper) increased the sustainability of the megasite, which, in exchange, provided a key context for inter-community ritual and exchange, as well as all of the other benefits arising in assembly places (Nebbia et al., 2018). It is suggested that the pre-existing links between the settlements of a site cluster, whether based upon the Big Other or exchange networks, would have been fundamental in the possibility of the emergence of a larger site serving all others in the site cluster and probably beyond—the region’s earliest megasites. This dynamic settlement system allowed the emergence of more than one megasite in a single cluster, indicating variations in the success of alliance-formation and an element of competition between these sites.10

Clearly, the Trypillia megasites did not stop in the BI Phase but continued for a further 600 years (4000–3400 BC). We shall content ourselves here with a summary of the major changes that took place in megasites in Phase BII11, using Nebelivka as an example. Although on the global CT level, Phase BII was marked by a fall in the number of sites, this was anything but the case in the Southern Bug—Dnieper Interfluve, where the number of site clusters grew to cover large parts of the network of smaller streams (Figure 5D). In this Phase, we can detect the emergence of the first megasites based in the smaller stream networks—sites such as Nebelivka. It is interesting to confirm that, despite the local increase in both settlement numbers and site sizes, the 100-km. interaction zone continued to operate for megasites such as Nebelivka. However, with the increase of settlement numbers, not only the size of megasites grew, but Phase CI sees the emergence of “isolated” megasites, such as Yaltushkiv I, Stina, Bilohorodka, and Obukhiv, that developed outside the Southern Bug—Dnieper interfluve, but that maintained the 100-km scale of interaction (Figure 11). This could have important implications on the meaning of the Southern Bug—Dnieper interfluve as the area of megasite emergence that progressively loses its place-value, during a time of Trypillia centrifugal expansion into new territories. This movement maintained the practice of megasite building and large-scale interaction for 200–300 years until their demise in Phase CII.

The most obvious differences between Nebelivka and the BI megasites concerns site planning and the appearance of a series of public buildings we have termed “Assembly Houses.” A greater degree of formalisation of planning is inherent in the integration of all four main planning principles in the Nebelivka plan (Figure 9). However, at the same time as the major planning elements have been strengthened as a consequence of bricolage, the size of the building project enabled local diversity in building design and location at all scales of the plan, from individual houses to Neighbourhoods (groups of houses), Quarters (groups of Neighbourhoods) and major planning elements (e.g., the variations in the width of the space between the Outer and Inner house circuits) (Chapman and Gaydarska, 2016). We have argued that local architectural diversity probably marks not only the contribution of many communities in the Nebelivka interaction zone to dwelling on the megasite but also the passage of social time in the creation of different built ensembles (Chapman and Gaydarska, 2018a).

---

10The same pattern of coeval megasites was to be seen in the early 4th millennium BC in the case of Nebelivka, Taljanki and Majdanetske (Millard, 2019).

11For a long-term account, see Gaydarska, 2019a, Chapter 6.
The apparently novel aspect of BII megasites concerns the creation of public buildings (“Assembly Houses”) to participate, if not take a lead, in local and trans-megasite ceremonies, including processions (Chapman and Gaydarska, 2019). Geophysical investigations at Nebelivka have produced the first and currently only complete megasite plan with modern geophysical instruments (Chapman et al., 2014a; Hale et al., 2017). These investigations have revealed the existence of 23 Assembly Houses, unevenly dispersed across the megasite but mostly outside the two house circuits. The location of the Assembly Houses was one of the criteria used to divide the megasite into Quarters (Chapman and Gaydarska, 2016) (Figure 9), producing a sense of a special local relationship between Neighbourhoods and “their” Assembly House. It is intriguing to note that the Assembly Houses were burnt in a quite different way from usual dwelling houses (Figure 12), reinforcing the difference between the two architectural forms. It is apparent that the building of Assembly Houses was one response to the much greater social and architectural complexity found in the BII megasites in comparison to their BI predecessors, contributing the increased formalisation seen in the larger BII sites.

The principal material culture changes from Phase BI to Phase BII concerned the decline in the quantity of lithic deposition, the increased deposition of painted pottery and the production of heavy copper tools. Greater reliance on local sources was probably one of the factors involved in the change in lithic deposition but changes in the operational chain were also involved. Two of the most significant effects of the innovation of Phase BII painted ware were the constant new demand for black, manganese-based pigment for potters in each community and the re-orientation of copper exchange toward Transylvanian sources. These changes led to a major expansion in inter-regional exchange, with high-quality lithics, copper and manganese pigment all brought from the Western part of the CT distribution to the Southern Bug—Dnieper interfluve. We are currently unaware of the linkage of the lithic, copper and pigment networks but they may have been closely integrated, with the same traders moving all three materials, at least East from the Prut valley.

How can these considerations be “translated” into an answer to the question of why the megasites emerged when they did, in Trypillia Phases BI and BII? There is no straightforward answer to this question, since we are dealing with a multivariate issue with many relevant data sets. The growth of settlement clusters in Phase BI led to increased interaction between the neighbouring settlements, which further increased in intensity with the need for buffering for the largest site in each cluster—the early megasites. The differential attraction of copper, lithics, and pigments of these early megasites helped to maintain their position as central assembly places in the face of their weakness—the absence of social mechanisms, perhaps principally planning mechanisms, to integrate visitors from large numbers of smaller settlements. This weakness in social controls would have led to either megasite abandonment or, as happened later, in Phase BII, to the emergence of planning practices which helped megasites to live more cohesively in even larger sites.

Another key aspect of Phase BII settlement in the Southern Bug—Dnieper interfluve was the increasing interaction between as well as within settlement clusters, which increased the value of co-ordinating assembly sites. The expansion of exchange to bring three critical resources—exotic flint, pigments for painted pottery and copper—from the same regions to the West further consolidated the BII megasites as assembly places for in turn larger settlement clusters. It should not, however, be forgotten that a megasite could fail at any time—there were many possible pathways to disintegration and decline. It is a mark of the stability of the social practice at the BII megasites such as Nebelivka that they continued for five or six generations before their ultimate demise.

**BUT WERE THE TRYPIILLA MEGASITES “CITIES”?**

In the etymological dictionary *Origins*, the term city is defined as an “aggregation of citizens” (Partridge, 1983, p. 101). As
clearly elucidated by Emberling (2003), this highlights three “basic” elements of the city, (1) a community of people with forms of social and political organization which are different from pre-urban and non-urban communities; (2) the aggregation happens in a specific location, the city, which is a physical space and a conceptual map of urban residents and their neighbours; (3) the inhabitants—citizens—identify themselves with the physical space, thus creating an urban identity (Emberling, 2003, p. 254). But what kind of urban identity?

We have already made a case that Trypillia megasites would not fit what we broadly call the “traditional” view of urbanism (Liverani, 2006; Gaydarska, 2016, 2017) and would be more at home with massive global phenomena still awaiting their name (“Big Anomalous” sites, “Big Weird” sites) (Fletcher, 2009). Some of these sites (e.g., Angkor) are the first to be recognized as low-density urban settlements (Fletcher et al., 2015), while Trypillia megasites are currently the earliest example of low-density occupation in well-defined large sites. We have also posited a relational approach whereby the meaning and function of given sites is only definable in relation to other sites, instead of in fixed and absolute terms (Gaydarska, 2016, 2017). In the CT context, that would replace the unhelpful site hierarchies based on size (Ellis, Videiko, Diachenko) and identify to what extent significant social practices differed from site to site. Ideally, such a comparison would involve settlement planning, depositional practices, subsistence practices, and the consumption of exotic and local objects made of clay, metal, and stone. Holistic inter-site evaluations are limited by more than 100 years of CT investigations, mostly based upon small-scale excavations and heavily biased toward pottery comparisons and classification. Still, there is some patchy evidence allowing the differentiation of sites and forms of human occupation. First, there is a tendency toward increasing settlement size, peaking in the 100 ha site of Kharkivka and the 150 ha site of Vesely Kut. Such social experiments would have accumulated practical experience of ways of mitigating the social tensions arising from scaled-up habitation. However, we know very little about the spatial arrangements at these early large sites. By contrast, other sites, such as Mogylina III, evince evolving principles of house concentricity among the more general pattern of a lack of formal planning but their size is very small (10 ha). The pattern in the Early Trypillia period (Phases A and BI) shows a contrast between some small sites with developed planning elements and other large sites with no evidence for evolved planning features. The proposed conclusion is that these two aspects of site development did not come together until Phase BII, at sites such as Nebelivka.

There are strong environmental indicators for human presence at the site of Nebelivka well before the establishment of the BII megasite but no material trace of such occupation has been found as yet. The implication is that short, probably temporary, but intensive and perhaps massive aggregations must have taken place that would account for both the strong human impact on the landscape and the lack of material evidence. Thus, although the “norm” for a Trypillian BII settlement was a small site with few distinctive planning elements and variable consumption of material culture, there were formalized and non-formalized forms of human occupation that deviated from that norm: settlements constituted the former, assembly places, and gathering places the latter. Taken individually and spread over some distance and in time, these differences may have not been perceived as “too different” and therefore threatening to the social order but remaining as part of the Big Other. But when ancestral memory and intensified human interaction in the BII period brought various practices together, this resulted in the creation of a very different kind of place—the 238 ha megasite of Nebelivka, with its intricate combination of formal layout and local diversity. In relational terms and according to the currently published data, the BII Nebelivka megasite stood out among its contemporary and preceding settlements. This was an emergent settlement form rooted in previous forms of dwelling and aggregation, whose novel combination marked a significant difference in relation to other sites. It was perceived, experienced and functioned as a very different kind of place that fulfilled a dual purpose of dwelling and assembly. It is in this sense that we see the megasites as what, in hindsight, modern scholars call “cities.”

CONCLUSIONS

The Trypillia megasites of the Southern Bug—Dnieper Interfluve in central Ukraine are the largest, and earliest, settlements in 4th millennium BC Eurasia and potentially the world; we claim that they are the earliest known cities. The megasites were not permanent, long-term settlements but have been modelled as different forms of low-density city, whether permanent with a much smaller population or as seasonal forms of assembly or pilgrimage places.

In this article, we propose a model for the origins of Trypillia megasites more consonant with this alternative view of smaller-scale settlements. Pre-existing exchange networks moving exotic flint, copper and salt across the forest steppe helped to consolidate the Trypillia Big Other as an ideological framework for building material traditions. Out of the mix of large, amorphous settlements and small sites with developed planning elements, but not both on any single site, emerged the BII megasites—an unprecedented settlement form where bricolage of earlier plan elements produced formalised sites which combined an inner assembly space with an outer dwelling space. Settlement modelling showed the scale of megasite interaction to remain stable at c.100 km for many centuries, integrating increasing numbers of small sites to megasite assembly places.

Because of their size and seasonality, Trypillia megasites benefited from the increasing connectivity of their 100-km networks and the specialised building of public buildings and production of painted pottery without suffering from the disadvantages of inequality, severe human impacts on the local landscape and lower standards of living. These developments enabled the reproduction of megasite lifeways for over 600 years, even though the lack of hierarchical structure prevented the appearance of successor settlements on the forest steppe.
AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

Grateful thanks to Francesca Fulminante for her kind invitation to contribute to this volume and for the OPEN-AIRE funding linked to the Marie Skłodowska Curie Project Past-people-nets 628818 conducted by Francesca Fulminante (2014–2016).

Deepest gratitude to Durham University and the Kyiv Institute of Archaeology, especially Dr. Mykhailo Videiko, for project support; thanks for financial support to the AHRC for Research Grant AH/1025867, National Geographic Society (Grant 2012/211) and the British Academy (S54204); thanks to the Kirovograd, Novoarkhangelk, and Nebelivka authorities for support; and thanks to the editors of this Issue and the reviewers who have improved this paper substantially with their comments. We are also grateful to Dr. Knut Rassmann for allowing us to reproduce his geophysical plot of Singerei as Figure 8.

REFERENCES

Bucure¸ sti: Muzeul National de Istorie a României.


Rassamakin, Y. (2012). “Absolute chronology of Ukrainian Tripolye settlements,” in The Tripolye Culture Giant-Settlements in Ukraine. Formation, Development...


von Stern, E. (1900). The meaning of ceramic finds from the South of Russia for the research of the cultural history of Black Sea colonisation. ZOOID (Berlin) 22, 1–21.


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Chapman, Gaydarska and Nebbia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Cities Through the Ages: One Thing or Many?

Michael E. Smith*1 and José Lobo2

1 School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, United States, 2 School of Sustainability, Arizona State University, Tempe, AZ, United States

The variability among cities, from the ancient world to the present, can be organized usefully in two ways. First, a focus on the dominant urban activities and processes leads to the recognition of two basic urban types: economic cities and political cities. Most cities today are economic cities in which growth proceeds through agglomeration processes. By contrast, most cities in the ancient world (and some today) are political cities, in which power and administration play a major role in structuring cities and generating change. Second, an alternative focus on processes of social interaction within the urban built environment leads to the recognition that there is only one kind of settlement that includes all cities—economic and political; past and present. Cities in this sense are settings for “energized crowding.” Processes of interaction generate both economic and political growth, and they produce and influence the built forms and social characteristics of all cities. Our model helps scholars distinguish the unique from the universal traits of cities today and in the past.

Keywords: cities, agglomeration, growth, comparative analysis, networks, scaling, social interactions

INTRODUCTION

Cities today are both the engines of innovation and economic growth, and the settings for concentrated social problems. As cities around the world expand in size and impact, advances in the scientific understanding of cities, urbanism, and urbanization take on increasing urgency. Do fundamental urban processes exist that generate a basic kind of city, in whatever context or culture or time period? Or are the expressions of urbanism too diverse to include in a single model? Is contemporary urbanization simply an elaboration upon past urbanization processes, or is it a fundamentally different kind of process?

Urbanization through the ages has manifested an enormous variation in the spatial and social forms of cities, their size, functions, activities, and growth patterns. Given this great variability in so many domains, it is not hard to argue that any notion of a single urban form or process or pattern throughout history must involve over-simplification. The differences among cities, across space and time, would seem too great to fit into a single type or model. Nevertheless, scholars in a variety of disciplines have argued that cities—regardless of their size, geography, temporal setting, or cultural milieu—share many underlying social characteristics, and play similar functional roles in different human societies (Mumford, 1961; Jacobs, 1969; Hall, 1998). There is a growing recognition that human settlements from Ur to Mumbai share enough in common that the term “cities” can be used to meaningfully refer to entities separated by thousands of years (Algaze, 2008; Smith, 2010b; York et al., 2011; Barthel and Isendahl, 2013). Yet this work has failed to specify just what characteristics of cities through the ages allow them to be considered or analyzed as a single phenomenon, in spite of their obvious differences.
Cities (and human settlements more generally) are essentially about the advantages afforded by aggregation which in turn are a manifestation of human sociality (Boyd and Richerson, 2005). Population size is both a major determinant and consequence of social evolution (Henrich, 2015). The basic demographic dynamics of expansion, maintenance, and decline are essential to cities, ancient, pre-modern, and modern. Over their existence all cities documented by archaeology and history went through phases of expansion (in spatial extension and population size) before eventually declining. The existence, extent, and relevance of economic growth (which can mean either increases in material output or increases in output per capita) in ancient and pre-modern societies is a topic of considerable debate (Erdkamp, 2016; Jongman, 2016; Stark et al., 2016),(Greene, 2000; Scheidel, 2004; Pryor, 2005). The reality of urban growth, on the other hand (either an increase in the proportion of a society's population residing in urban settlements or the increase in the population size of individual urban settlements), in ancient and pre-modern societies is not. Understanding what is common and what is not with regards to urban (population) growth across time is key to understanding what is common to urban life across eras and civilizations.

The question posed in our title—one thing or many?—has two answers. First, when we focus on the institutional framework of cities, including the mechanisms that generate urban growth, there are two fundamentally different forms of cities. We call these economic cities and political cities. In urban economics, this contrast is often discussed in terms of differences between “normal” cities and primate cities (de Long and Shleifer, 1993; Ades and Glaeser, 1995), but in fact the distinction runs deeper than this. Most cities before the modern era were political cities, meaning that their dominant institutions were in the realms of power and administration by a ruler or ruling elite. A few pre-modern cities and most cities today are economic cities, meaning that economic considerations dominate the locational decisions of individuals and production units, and that economic activities largely shape their social structure and economic forces dominate their processes of growth.

Notwithstanding the very real and important differences between economic cities and political cities, urban growth in both types originates in a common set of behavioral and built-environmental mechanisms that underlie both the economic and political drivers normally discussed in the urban literature. A focus on these fundamental mechanisms leads to our second answer to the question of “one thing or many?: the city is one thing amidst a plurality of manifestations of urban life. It is a place of energized crowding (Kostof, 1991, p. 37) that generates growth and change (Smith, 2019). These underlying mechanisms have been explored by several strands of recent work that take the perspective—-theoretically grounded and empirically supported—that cities are, and have been, social networks of people embedded in physical space (Fisher, 2009; Hipp et al., 2012; Bettencourt, 2013; Youn et al., 2016). This body of research has revived the prospects for building an analytical framework for understanding the origins and drivers of urbanization operating in cities from the distant past and contemporary urban life, both economic, and political cities.

Our dual answer to the question of “one thing or many?” has implications for understanding cities and urbanism today. Whether one is interested in identifying universal urban traits (Smith et al., 2015), tracing the development of urbanism over time (Mumford, 1961), or using ideas from past cities to inform contemporary practice (Rapoport, 1973; Hakim, 2012), it is important to understand both the continuities and disjunctions between cities today and those of the past. Are the principles of the new urbanism (Congress for the New Urbanism, 1996) based on universal urban realities, or do they only reflect conditions of the very recent past? How can scholars predict whether current principles of urban resilience will play out over long periods of time? We propose that our formulation of cities through the ages can help frame analyses of these and other questions about cities and urban life in the present and future.

WHAT IS A CITY?

In his book Triumph of the City, urban economist Edward Glaeser defines cities as “the absence of physical space between people and companies. They are proximity, density, closeness” (Glaeser, 2011, p. 6). Glaeser's minimalistic definition does seem to capture what for many is the essential feature of cities. Yet its simplicity eliminates the definition to apply equally to every form of human settlement—from hunter-gather camp to city—in which physical proximity facilities social life. Furthermore, high density is no longer characteristic of all modern cities (Angel, 2012), nor was it characteristic all ancient cities (Fletcher, 2009).

Archaeologist Cowgill (2004, p. 526) observed that, “It is notoriously difficult to agree on a cross-culturally applicable definition of the city, but we cannot do without definitions altogether... No single criterion, such as sheer size or use of writing, is adequate.” The urban literature reveals two dominant approaches to city definition: a sociological/demographic approach and a functional approach. The most influential definition is that offered by sociologist Wirth (1938, p. 8): “For sociological purposes a city may be defined as a relatively large, dense, and permanent settlement of socially heterogeneous individuals.” This definition clearly fits contemporary cities, and it is favored by most scholars of urbanism today. But historians, anthropologists, and archaeologists have pointed out that the sociological definition excludes most pre-modern cities from consideration as urban settlements (e.g., Sjoberg, 1960; Fox, 1977; Smith, 2016).

Early urban settlements were prominent and influential within their regional settings, yet their levels of population size, density, and heterogeneity were considerably lower than contemporary western cities. This situation led to the adoption of a “functional” approach to definition, based on work in economic geography (Lloyd and Dicken, 1972). The definition of archaeologist Trigger (1972, p. 577) is typical: “It is generally agreed that whatever else a city may be it is a unit of settlement which performs specialized functions in relationship to a broad hinterland.” An urban function is an activity or institution that directly affects life and society in a hinterland. The initial
THE CITY AS TWO THINGS: ECONOMIC AND POLITICAL CITIES

Most cities before the Industrial Revolution were political cities and most contemporary cities are economic cities. This dichotomy is based on the nature of economic activity, its prominence in urban dynamics, and its relationship to growth. Virtually all of the literature in urban economics and urban geography focuses on the locational choice of businesses and individuals, and on the positive externalities (due to agglomeration) to the productivity of firms and individuals. This leads contemporary work in urban economics, economic geography, and regional science to have a “modernist” bent applicable mainly to cities whose economies are productive and dynamic. In our scheme most economic cities exist within the capitalist world system, but we think that more research is needed to determine the extent to which our economic cities concept might apply to cities before the modern era. Political cities, on the other hand, are those cities in which political or administrative activities predominate. Economic processes are either suppressed by political forces (as in recent primate cities), or else are simply far less developed than in contemporary cities.

The concepts economic city and political city are Weberian ideal types: pure classifications that will never match the empirical world precisely (Gerth and Mills, 1946, p. 59, 60). They are ends of a continuum. This dichotomy is similar—but not identical—to a number of dichotomous typologies in the urban literature (Table 1). We are proposing a new dichotomy because none of these prior schemes are sufficiently broad to encompass the entire range of historical urbanism.

The earliest of these ideal-type city dichotomies was Smith (1979) distinction between government towns and commercial towns (Stull, 1986). The primary activities in the former are administration and rule, which, in Smith’s scheme, are considered unproductive labor. Smith describes these workers as “idle, dissolute, and poor” (Stull, 1986, p. 300). Most workers in commercial towns, in contrast, are “industrious, sober, and thriving.” Adam Smith’s list of eighteenth-century government towns includes Rome, Madrid, Versailles, Paris, and Edinburgh prior to 1707. Commercial towns included Glasgow, many English towns, and most Dutch towns. He classified other cities of his time as having attributes of both types: London, Lisbon, Copenhagen, and Edinburgh after 1707.

The contrast between the consumer city and producer city originated with Max Weber and other early twentieth century economic historians to contrast the ways in which Classical and medieval cities obtained food from their hinterland. Commercial enterprises in medieval producer cities allowed urbanites to obtain food from farmers through commercial exchange, whereas elites in ancient consumer cities received rural income from rents. Classicists debated which concept best fit Roman cities and towns for decades (e.g., Parkins, 1997), an argument that got tangled up with the primitivist/modernist debate on the ancient economy. Recent work has reached two conclusions: (1) markets were operational in many but not all ancient societies; and, (2) the presence of markets does not turn ancient societies

---

TABLE 1 | The city as two things: dichotomous city typologies.

<table>
<thead>
<tr>
<th>Political cities</th>
<th>Economic cities</th>
<th>Citations</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government towns</td>
<td>Commercial towns</td>
<td>Smith, 1979</td>
<td>Cities in Wealth of Nations</td>
</tr>
<tr>
<td>Consumer city</td>
<td>Producer city</td>
<td>Weber, 1958; Finley, 1973</td>
<td>Classical vs. medieval cities</td>
</tr>
<tr>
<td>Public urbanism</td>
<td>Commercial urbanism</td>
<td>Clarke, 1993</td>
<td>Roman cities in Britain</td>
</tr>
<tr>
<td>Parasitic city</td>
<td>Generative city</td>
<td>Hoselitz, 1955</td>
<td>Cities in developing nations</td>
</tr>
<tr>
<td>Orthogenetic city</td>
<td>Heterogenetic city</td>
<td>Redfield and Singer, 1954</td>
<td>Cultural roles of cities</td>
</tr>
<tr>
<td>Dependent economy</td>
<td>Autonomous economy</td>
<td>Fox, 1977</td>
<td>World historical cities</td>
</tr>
<tr>
<td>Primate city</td>
<td>Non-primate city</td>
<td>Various</td>
<td>Deviation from Zipf’s law</td>
</tr>
</tbody>
</table>

---

1We realize that some scholars in the humanities view typologies—particularly dichotomies—as simplistic devices that impede understanding of the rich details of individual cases. But comparative analysis and theoretical advance require simplification in order to promote understanding on a level above that of individual cases (Healy, 2017; Smith, 2018). Simplified schemes—like economic and political cities—are not intended to substitute for detailed studies of specific cases or contexts; instead, their purpose is to promote understanding and explanation. They complement detailed studies of cases.
into protocapitalists entities (Morris, 2004; Feinman and Garry, 2010; Garry and Stark, 2010).

In a study of variation in Roman cities in Britain, Clarke (1993) evaluates several general models of the spatial organization of preindustrial cities, including Sjoberg (1960), Vance (1971), and Langton (1975). None of these models fit the Roman cities of Britain well, leading Clarke to synthesize various insights of these authors into two models that he calls, “Public urbanism” (Sjoberg’s model of class-based spatial zones, with less emphasis on the commercial component) and “Commercial urbanism” (the Medieval city model of Vance, which emphasizes the importance of guilds and the lack of zoning by class).

The next two dichotomies arose in the 1950s in the literature on economic development. Economist Hoselitz (1955) classified cities as generative or parasitic if they have a positive or negative impact on economic growth in their region or country; see also Wrigley (1978). Anthropologists Redfield and Singer (1954) promoted stereotypes of developing nations in their classification of orthogenetic cities (traditional cities where the “moral order” dominates) and heterogenetic cities (modernizing cities where the “technical order” is primary). The dependent/autonomous economy dimension of Richard Fox’s functional typology maps onto our political/economic dichotomy quite closely; his regal-rural, administrative, and colonial cities categories fit into our political category, and his industrial and mercantile cities fit into the economic type. City-state capitals are split between the two categories; Greek, Yoruba, and Aztec examples were political cities, but many post-medieval European city-state capitals were economic cities. We also include primate cities in our political category. These are cities whose size far exceeds other cities in their regional system (Adamic, 2011). Of contemporary primate cities, [Ades and Glaeser, 1995], p. 195] state, “political forces, even more than economic factors, drive urban centralization” (we discuss primate cities more fully below).

Our categories political and economic cities—like the other parallel dichotomies listed in Table 1—are Weberian ideal types that are not intended to match precisely any specific city. In fact, these categories may best be viewed as ends of a continuum rather than as a rigid dichotomy. Our purpose is to improve scholarly understanding of the variability among cities, not to categorize specific cities.

Growth in Cities
The phenomenon of growth and its drivers would seem to provide the starkest difference between ancient and modern cities. Clearly delineating these is an important part of an exercise aiming to identify commonalities across the historical experiences of urban development. We start with the question, what is economic growth? At its most basic, an increase in a society’s material output with respect to a previous period is an instance of economic growth. An increase in economic output caused by more efficient use of inputs is referred to as intensive growth, while growth caused only by increases in the “factors of production” (such as labor or agricultural land) is called extensive growth (Bjork, 1999). Labor power was and remains the most important input into production, and therefore population growth alone would have sufficed, in most situations, to bring about an increase in material output in pre-modern societies (otherwise a process of immiseration would have kicked in.) “Smithian growth” refers to a situation in which growth is driven by increased labor specialization, itself facilitated by the geographical expansion of markets (Burkai, 1969; Kelly, 1997; Persson and Sharp, 2015). This type of growth required an extension of commercialization and transportation infrastructure. The Solow–Swan type models of economic growth, with their explanatory emphasis on capital accumulation and population growth, exerted great influence in the economics profession during the 1960s and 1970s (Solow, 1956; Swan, 1956). In the modern era economic growth, of the Smithian or intensive variety, has entailed increases in productivity (usually captured via measures of output per capita) so much so that—for modern economies—economic growth is tantamount to increases in productivity (Allen, 2009).

For nearly three decades now work in growth economics has been dominated by the “new economic growth theory” which emphasizes the generation and exchange of knowledge, innovation and invention, and human capital (i.e., skilled individuals) as drivers of productivity increases and growth (Lucas, 1988; Romer, 1990; Weil, 2008). The emphasis on knowledge spillovers (a form of externality) in turn gave rise to a renewed interest in the role of cities as the privileged setting for the generation and recombination of knowledge (Lucas, 1988; Glaeser, 2011). The manifestation of the “new growth theory” perspective with respect to urban development has emphasized the operation of “agglomeration economies” (Fujita et al., 1999). Agglomeration economies arise when concentrations (“agglomerations”) of individuals, firms, and institutions create interactions and feedback that generate non-market mediated benefits (knowledge flows, for example). These agglomeration economies are assumed to be major drivers of growth in most recent works explaining urban economic development (e.g., Black and Henderson, 1999; O’Sullivan, 2011; Storper, 2013).

Duranton and Puga (2004) divide the forces generating agglomeration into three micro-level processes: sharing, matching, and learning. Sharing refers to the presence of public goods (infrastructure, markets, and other institutions facilitating commerce) in cities, and to the gains from specialization and from larger number of suppliers that are shared among individuals and firms. Although political cities certainly offered public goods (Stanley et al., 2016), the much lower levels of technology, infrastructure, and commercialization led to lower levels of productive interaction and economic growth. Matching refers to the pairing of people and jobs that occurs in urban areas. Given the far lower prevalence of wage labor in pre-capitalist economies—coupled with the far lower level of individual specialization—it is unlikely that matching was a significant force in creating agglomeration in political cities. Learning refers to the generation, diffusion and accumulation of knowledge, particularly the education of workers that contributes to human capital in agglomeration economies. The much reduced levels of education and literacy in pre-capitalist economies renders learning a minor factor in generating change and economic growth in political cities. Storper (2013) provides a parallel discussion of the major causes of urban
growth and agglomeration. From a broader perspective, if we consider the realm of contemporary economic growth—not just urban agglomeration—it is quite clear that most of the major processes (Jones and Romer, 2010) have few counterparts in pre-capitalist economies.

**Political Cities and Economic Growth in the Ancient World**

The concept of primate city in urban economics (de Long and Shleifer, 1993; Ades and Glaeser, 1995; Behrens and Bala, 2013) gets to the core of the distinction between the growth processes of political and economic cities. It is in the realm of economic growth, relentlessly present in modern social life and seemingly absent in the ancient past, that the greatest chasm between ancient and contemporary urban life is to be found. [Ades and Glaeser (1995), p. 224] summarize how primate cities differ from other cities as follows:

Urban giants ultimately stem from the concentration of power in the hands of a small cadre of agents living in the capital. This power allows the leaders to extract wealth out of the hinterland and distribute it in the capital. Migrants come to the city because of the demand created by the concentration of wealth, the desire to influence leadership, the transfers given by the leadership to quell local unrest, and the safety of the capital. This pattern was true in Rome, 50 C.E., and this description is still true in many countries today.

[de Long and Shleifer (1993) 1993, p. 686] note for their historical sample that, “the presence of an absolutist prince [a marker of a primate city] reduces the growth of population in cities of more than 30,000 by nearly 180,000 people per century.”

The level of commercialization in an economy provides a rough index of the scale from political to economic cities. By “commercialization” we mean the number and influence of commercial institutions in an economy. Such an index runs from uncommercialized economies that lack money, accounting systems and entrepreneurial merchants, to moderately commercialized economies (with money in the form of coinage and accounting systems, but without wage labor or banking), to highly commercialized economies (with all of these traits and others). This concept and scale is based on Smith (2004) and a series of works in economic history (e.g., Braudel, 1982; North, 1991; Temin, 2013; Persson and Sharp, 2015).²

The remarkable thing about such a this scale of commercialization is that even the most highly commercialized pre-Medieval economy—Imperial Rome—still had very limited levels of economic growth. Growth did occur in ancient Greece (Morris, 2004) and Imperial Rome (Jongman, 2012), but Roman cities were not dynamic and expanding commercial centers in the way that contemporary economic cities are. There was a real transformation between the Roman Empire and the medieval period in the nature of commercial activity and the role of cities.

Double-entry bookkeeping, for example, did not begin its rapid expansion until around 1,500 (Gleeson-White, 2012).

Hudson (2010) pinpoints the activities, institutions, and values of the Classical world responsible for what he calls the “corrosive forms of enterprise” that dominated the economy. Rent-seeking behavior was rampant. “The oligarchic ethic preferred seizing wealth abroad to creating it at home. The major ways to make fortunes were by conquest, raiding and piracy, slave capture and slave dealing, money lending, tax farming, and kindred activities more predatory than entrepreneurial” (p. 15). Although the economy of Imperial Rome had reached the highest level of commercialization in the world prior to Medieval Europe, its economy was still “unproductive” (Baumol, 1990) in the prevalence of rent-seeking over commercial growth. In a study of Roman banking, [Andreau (1999), p. 147–48] asks, “Did Roman financiers direct most of their effort toward economic life in order to create an effective instrument for investments? Did any financial establishments specialize in the promotion of productive loans? The answer to both questions must definitely be no.”

Of the Roman Empire, [Hopkins (1978), p. 77] noted, “Huge pre-industrial empires accumulate huge resources; they spend a large part of that accumulated surplus on self-preservation, not on economic growth.” The implication is clear: the behaviors and institutions that produce economic growth and urban agglomeration economies in modern cities were for the most part absent from the political cities of the ancient world, and these cities require very different models of growth. In the words of historian [Noreña (2014), p. 193] “All cities in preindustrial economies are in one sense artificial, in that they depend on a set of political institutions, coercive instruments, and legitimizing mechanisms that together enable a group of non-primary producers to live off the surplus produced by peasant farmers.”

**Agglomeration Processes in Political Cities**

Even if we acknowledge that economic growth was far less important in ancient cities than it is today, it is still useful to ask whether agglomeration economies might have generated urban growth in the political cities of the past. Virtually the entire literature on urban agglomeration is tailored for contemporary or recent historical economic cities with capitalist economies. Nevertheless, it is possible to expand the concept of urban agglomeration beyond economics to make it more applicable to pre-modern cities. This is the approach taken by several economists to explain anomalous patterns of growth in historical and contemporary primate cities (de Long and Shleifer, 1993; Ades and Glaeser, 1995). Their approach can be broadened even further by considering the spatial concentration of various kinds of non-economic activity in pre-modern cities and its effects.

Urban agglomerations of power and administration are the most obvious case, but it is not too far-fetched to suggest that ritual activity could also generate concentrations of social interactions and urban growth. Indeed, this was one component of Wheatley’s model of early urbanization. In a recent paper, [Scott and Storper (2015), p. 4] propose this kind of broader approach to agglomeration:

---

² A commercialization scale was devised for a project that compared 23 premodern cities on a variety of attributes (Smith et al., 2016; Stanley et al., 2016), and I thank Barbara Stark, Benjamin Stanley, Abigail York, and Timothy Dennehy for help in constructing this commercialization scale.
Even in the very earliest cities, agglomerations of activities such as political administration, ceremonial and religious pursuits, craft production (e.g., for luxury goods or military hardware), and market trading almost always constituted the core of the urban process (Wheatley, 1971). Agglomeration occurs because activities like these entail divisions of labor and other interdependencies as expressed in transactional relationships whose costs are distance dependent and because they can reap functional synergies by clustering together in geographic space. Various types of infrastructure help to consolidate the resulting dynamic process of agglomeration. In other words, one of the central features of urbanization has always been its efficiency-generating qualities via agglomeration.

Nevertheless, the levels of feedback and urban growth generated by these non-economic agglomeration processes seem orders of magnitude smaller than the contemporary growth processes of agglomeration in economic cities.

Scott and Storper seem to recognize the gulf between agglomeration processes in political and economic cities when they point out that early cities were “caught in a Malthusian trap” that was only overcome with the sharp rise in economic productivity with industrialization. They suggest that the Industrial Revolution was “an era where the fundamental relationship between economic development and urbanization becomes especially clear” (Scott and Storper, 2015, p. 5). But from our perspective, this latter statement is incorrect. Industrialization did not clarify the relationship between economic growth and urbanization; instead, industrialization, and capitalism both transformed and greatly strengthened this relationship to the point where the dynamics became fundamentally different from the growth processes of political cities, in the sense that “more is different” (Anderson, 1972). Nevertheless, we suggest that further attention to non-economic agglomeration processes can not only help explain urban growth in the past, but may also shed light on the variety of urban growth trajectories in the contemporary world.

**THE CITY AS ONE THING: A CONTAINER FOR “ENERGIZED CROWDING”**

The concept of cities as places of intensive social interaction—“energized crowding” in the words of [Kostof (1991), p. 37]—is shared by both of the dominant definitions of cities and urbanism (Smith, 2019). Furthermore, processes of social interaction operate at a deeper or more fundamental level than the economic and political drivers of urban growth reviewed above, and they can be considered as generative forces for the growth of both economic and political cities. Our argument draws on a wide range of social-science scholarship on this topic.

Population size, density and heterogeneity have long been recognized as fundamental properties of urban settlements (Wirth, 1938; Bairoch, 1988; Angel et al., 2016). This socioeconomic and cultural heterogeneity in turn facilitates the plethora of interactions (intentional or serendipitous, fleeting, and consequential, anonymous or long-lasting, driven by economic imperatives or sustained by shared ideological commitments) that make urban life sociologically distinct (Jacobs, 1961; Fischer, 1975). It is this crowding—and the creative, inventive and innovative possibilities which it realizes—that in turn generates growth and change (Glaeser, 2011). Storper and [Storper and Venables (2004), p. 31] have analyzed the mechanism through which “energized crowding” occurs. They call this, “the most fundamental aspect of proximity: face-to-face contact.” Although their analysis of the role of face-to-face interaction in urban agglomeration is set firmly within the economic geography approach that ignores political cities and pre-modern societies, at least three of their four “basic functions” of face-to-face contact apply to pre-modern societies: communication technology; trust and incentives in relationships; and screening and socializing. Their fourth function—“rush and motivation”—is more difficult to evaluate for pre-modern contexts.

The perspective that all cities—across eras, geographies, and cultures—share a core of fundamental socioeconomic processes as well as certain predictable quantitative properties has recently coalesced into settlement scaling theory (Bettencourt et al., 2007; Bettencourt, 2013; Bettencourt and Lobo, 2016). Settlement scaling theory draws on insights from urban economics, economic geography, and regional science and shares with these disciplines a common explanation for the existence and development of cities as resulting from the interplay between centripetal and centrifugal “forces” (Colby, 1933; Isard, 1956; Fujita et al., 1999; O’Flaherty, 2005). Population size is arguably among the most important determinants and consequences of socioeconomic development and change in societies before the modern era (Carneiro, 2000; Johnson and Earle, 2000). Such relations are known across the sciences as scaling relations, which relate the macroscopic properties of a system—here a city—to its scale, or size (Barenblatt, 1996, 2003; Brock, 1999). For this reason, the systematic study of such relationships in cities is known as urban scaling or settlement scaling.

What is novel about the scaling framework is that it views cities as integrated socio-economic networks of interactions embedded in physical space, and then derives specific quantitative predictions about the relationship between population size, material output and aeral extent of settlements. The social interactions facilitated by physical proximity—and the lower costs associated with such interactions—drive productivity (Bettencourt et al., 2008, 2013; Bettencourt, 2013). These relationship have been statistically investigated not only for modern metropolitan systems (in the United States, Western Europe, Brazil, Japan, India, China and South Africa) but also for Native American farming villages in North America (before the arrival of the Europeans), Pre-Hispanic Andean and Central Mexican settlements, Ancient Greek and Roman cities, Medieval European cities and towns, and cities in Tudor England (Ortman et al., 2014, 2015, 2016; Cesaretti et al., 2016; Hanson and Ortmann, 2017; Ortmann and Coffey, 2017; Ossa et al., 2017; Cesaretti et al., under review). The results show a striking similarity in scaling relationships, as predicted by the settlement...
scaling framework. Across cultures and history, when human societies create permanent settlements, these grow denser, on average, with growing population. In addition, the increases in economic production and outputs are proportionally greater than the increase in population size. This, in turn, indicates that humans in larger settlements live in sociocological settings that facilitate higher rates of social interaction, relative to smaller settlements in a given context.

Social settings of ongoing face-to-face contact generate processes of trust, incentives, and monitoring that are among the key ingredients of successful community organization (Ostrom, 1990; Sampson, 1999; Bowles and Gintis, 2002). Urban planners are much concerned with finding ways that the urban built environment can foster productive face-to-face interaction (Talen, 1999; Brower, 2011). The observation that neighborhoods are universal features—not just of cities, but of almost all types of large human settlements (Smith, 2010a; Smith et al., 2015)—supports the notion that face-to-face interaction is fundamental and important in all kinds of cities, both economic and political.

The city as a facilitator of social learning might ultimately provide one of the most salient sources of continuity among urban forms across space and time. After all, social learning is emblematic of Homo sapiens. The resulting process of cultural adaptation that is responsible for our species’ success was facilitated by both population size and social connectedness (Boyd and Richerson, 2005; Henrich, 2015). Urban life is just one manifestation of the way humans have constructed social and physical spaces to exploit their unique abilities to build cumulative culture (Enquist et al., 2008). We should note that not all of the outcomes of energized crowding are positive traits; increased poverty, crime, and social alienation are also products of social interactions in cities.

These observations about the generative role of face-to-face interaction, from a spectrum of academic disciplines, apply equally well to political and economic cities. The urban built environment provides settings for social interaction, whether the parks, cafes, and sidewalks of contemporary cities or the plazas and marketplaces of ancient cities (Stanley et al., 2012). The same processes of communication and exchange take place among the residents of all cities, past and present. These basic interactions are the essence of what is distinctive about cities as human settlements. Their expression in the broad similarities of scaling relationships across time and history has a clear implication: the city as a container for social interactions throughout history and around the world is one thing, not many.

**DISCUSSION AND CONCLUSIONS**

Our analysis has implications for identifying more and less productive paths of comparative research of cities and urbanization. The “city as two things” perspective suggests that analyses of urban growth in economic and political cities are best carried out independently because the processes of growth in the two types of city are quite different. Empirical analyses of the rates and mechanisms of growth in past cities can help identify economic cities before the modern epoch. We argue above that economic growth in Imperial Rome was insufficient to generate the agglomeration economies of modern economic cities. But, given the high level of documentation for Roman urbanism, quantitative analyses of Roman urbanism (e.g., Bowman and Wilson, 2011; Jongman, 2016; Hanson and Ortmann, 2017; Hanson et al., 2017) provide models that will contribute to a better understanding of the nature of political and economic cities in the ancient world more generally.

Attempts to apply models of growth devised for contemporary economic cities to ancient cities in the absence of extensive quantitative data (e.g., Algaze, 2008, 2018) may be premature. On the other hand, models of modern primate (political) cities (de Long and Shleifer, 1993; Ades and Glaeser, 1995) would appear to be a productive source of insights for understanding ancient political cities. And given the smaller number of such political cities today and their lack of prominence in the literature on growth theory, perhaps the growing research on urban dynamics in pre-modern political cities might help scholars better understand the primate cities that dominated many developing nations today.

Our second perspective—the city as one thing—has great potential for comparative research that draws cases from, and produces insights for, both pre-modern and contemporary cities. For example scholars are beginning to draw parallels between the roles of neighborhoods in ancient and modern cities (Smith, 2010a; Arnauld et al., 2012; Sampson, 2012; Smith et al., 2015). Our analysis suggests that this topic is a particularly fertile one for comparative analysis, given the importance of social interaction and the built environment in neighborhood dynamics. The success of urban scaling research in identifying the predicted scale relationships in ancient settlement systems (see the sources cited above) provides strong empirical support for this perspective.

Productive research comparing pre-modern and contemporary cities needs to avoid two approaches that have long dominated comparisons of ancient and modern social and economic conditions. On the one hand, many social scientists outside of anthropology simply assume that conditions in the past or in non-western societies were similar to conditions today, and thus economic models can be applied directly without modification. This approach is called “modernism” in Classical studies and “formalism” in economic anthropology. Such work tends to be ignored or dismissed by historians and archaeologists because much of it fits the facts so poorly as to be useless. On the other hand, some scholars (“primitivists” in Classics and “substantivists” in economic anthropology) have portrayed the past (and non-western societies) as so radically different from the present that comparisons are impossible (Polanyi et al., 1957; Finley, 1973). Most scholars now see this as a misguided and highly limiting approach to past societies and economies (Smith, 2004; Wilk and Cliggett, 2007).

Our dual answer to the question posed in this paper’s title (“one thing or many”) can help scholars avoid these problems of misleading comparisons. From the perspective of the forces that create cities and generate urban growth, there are two general types of cities: economic and political. The dynamics of growth and operation of these cities are quite different, and facile comparisons between them will do little to illuminate general processes of urbanization. But from the perspective of the way
that people interact with one another in delimitled urban spaces, all cities are similar; there is only one type of city. The “energized crowding” first identified by [Kostof (1991), p, 37] is one of the universal features of cities, from ancient time to the present (Smith, 2019). The ability of scholars to explain and understand this process of energized crowding may help determine the productivity of comparative urban scholarship in the future, and it may even contribute to the success or failure of cities and urbanization in the future.

AUTHOR CONTRIBUTIONS

JL and MS conducted the research and wrote the article.

ACKNOWLEDGMENTS

This paper originated in a public lecture by Smith at the Santa Fe Institute in July, 2013, and the ideas were subsequently developed through conversations and face-to-face interactions. We want to thank Scott Ortman and Luis Bettencourt for many productive discussions of cities and urbanization processes. We thank the Santa Fe Institute and the ASU-SFI Center for Biosocial Complex Systems for funding the Working Group in Urban Scaling, which provided the opportunity for us to develop these ideas. Smith wants to acknowledge the always stimulating face-to-face interaction with his colleagues in the research project, Urbanization through the Ages: Neighborhoods, Open Spaces, and Urban Life in the School of Human Evolution & Social Change at ASU. Breandan Ó’Hallacháin helped Smith wade into the murky waters of agglomeration economies with discussions and bibliographic suggestions. We thank Scott Ortman, Barbara Stark, and Abigail York for helpful comments on an earlier draft of this paper. Presentations at the University of British Columbia and Dartmouth College contributed to the refinement of our ideas.

REFERENCES

More Real Than Ideal: Household and Community Diversity at Metapontum, South Italy

Andrew H. F. Cabaniss*

Interdepartmental Program in Classical Art and Archaeology, University of Michigan, Ann Arbor, MI, United States

Empirical studies of ancient cities must break down communities into their component parts, but frequently encounter difficulty with the scarcity of excavated domestic structures (e.g., Kramer, 1982, p. 673). I introduce to the archaeological literature the entropy estimating statistical bootstrap (EESB), a tool developed in information theory and computational social science by DeDeo et al. (2013) which provides a way to assess how representative a small dataset is of a parent population, categorized according to some useful typology. This method can be used to decide when small datasets can add further detail to our quantitative studies of archaeological settlements or when they need to be rejected as too small. I then illustrate its uses within the context of urban demography by examining the distribution of house forms to calculate household characteristics specific to Metapontum, an ancient Greek city. Future applications will include building larger urban datasets that are empirically grounded in the specific evidence for each community, facilitating the work of research programs such as urban scaling.

Keywords: urbanism, diversity, entropy, households, ancient Greece, communities, metapontum

INTRODUCTION

To make sense of urbanism requires a study of communities and of small groups like households. Discussions of ancient urbanism in the classical Mediterranean frequently focus on institutional macroscales of analysis to the detriment of individuals, partly because well-published evidence for households and burials are unevenly distributed. Their scant traces are frequently difficult to interpret in isolation without understanding their relation to the absent majority of material remains. Making rigorous use of small datasets of houses is required to grapple with the full richness of the archaeological traces of ancient cities.

In this paper, I examine the rural domestic remains of the city of Metapontum in Southern Italy in order to explore its urban demography and social diversity. In part, this is because thousands of farms are known from the well-documented countryside through surface survey, yet only a handful of these have been excavated (Carter, 2011). In order to check whether the small sample of excavated farmsteads is consistent with the less well-known survey sites, I introduce into the archaeological literature the entropy estimating statistical bootstrap (EESB) (DeDeo et al., 2013). The EESB can be used to assess whether a categorized sample is consistent with a similarly categorized source population. I provide an example to clarify how it works, and then provide a toy model of houses in a fictitious community that would be rejected by this method because of an overly small sample size for too diverse a community when described by a particular typology.
For Metapontum, I use the EESB to show that the sample of 11 rural houses from the 6th through 3rd centuries BCE appears to be consistent with the inferred frequency of different types of houses present in the ancient city when the types are defined in a socially meaningful way. In short, passing the EESB test makes a demographic analysis of households at Metapontum feasible, despite the small sample size. I then estimate the average characteristics of these households and an overall mean household size for the city. Statistical tests and calculations based on small samples can extend current methods for the quantitative study of ancient urbanism by providing a more detailed empirical record unique to individual cities such as Metapontum. It is hoped that this method will aid studies that attempt to put ancient Mediterranean urbanism into quantitative comparative frameworks like urban scaling (Hanson and Ortman, 2017). The EESB enables urban scaling theory to proceed at the level of populations as it was originally framed (Bettencourt, 2013), and in general gives the diversity of ancient cities the chance to speak for themselves within large-scale studies.

Houses and Families at Metapontum

The community at Metapontum, an ancient Greek city in the south of the Italian peninsula, has an extensively documented material record because of the density of excavations and surveys throughout the city's hinterland. The record of rural housing provides an opportunity to study a classical Mediterranean urban community and assess whether the published houses offer a representative glimpse into the demographic diversity of its inhabitants.

Much as in modern cities where the urban population is defined based on commuting patterns and social interactions (Bettencourt et al., 2010), so in many ancient Greek cities city folk had country houses in which they lived full- or part-time while engaging with the political, economic, religious, and social life of the city. Carter has argued that at Metapontum civic infrastructure was explicitly expanded to enfranchise the rural inhabitants and include them in political decision making (Carter, 2006, p. 204–24). It is important to remember these rural dwellings in our demography, and in the case of Metapontum they are more fully published, making it easier to consider how house architecture reflects household priorities, which in turn suggest demographic differences between households. While the presence and documentation of an urban grid have influenced our view of Metapontum just like many ancient Greek cities, there is also considerable evidence for the importance of the countryside as a residential option for the urban population.

History of Metapontum

Metapontum, or Metaponton, emerged in the eighth century BCE through violent conflict, persistent trade, and other intense modes of interaction at the intersection between indigenous, colonial, and trading populations (Carter, 1980; De Siena, 2001). Over a half century, the ancient Greek community disrupted the existing way of life for coastal and foothill Italic populations through aggressive raiding and territorial expansion, causing the abandonment of settlements in and around the future site of the city (Carter, 1980, 2006, p. 197–203; De Siena, 2001). Networks of sanctuaries and villages expanded into the countryside during the 6th century as the city center itself first defined its form with an urban grid, communal architecture, and civic temples (Osanna, 1992, p. 45–46; Carter and Prieto, 2011, p. 559–67; Carter, 2006, p. 204–14).

The Archaic boom of the city preceded an uneasy period of growth and contraction. Early fifth century land reclamation projects encouraged an expansion by farmers onto new lands, but a half century later these farms lay abandoned and new civic constructions remained without repairs for years (Carter, 2006, p. 214–24; Mertens, 1999). Individuals likely moved back within the city walls full time, or alternatively left the area entirely because of neighboring political instability. The Lucanians who began to move into this area may have responded to or caused some of these disruptions, but ultimately benefited from a city eager to have its countryside resettled (Carter, 2006, p. 218–32).

Fourth century Metapontum’s investments in the future of the town did not provide stability in an increasingly Roman Italy; following their alliance with Pyrrhus against Rome, a military camp was imposed on the city, at which point the population began to emigrate en masse (De Siena and Giardino, 2001). Opposing the Romans during the Second Punic War led to the city being carved up into farmland for massive estates that would come to dominate the Roman Italian agricultural system.

The community at Metaponto thus ebbed and flowed into its countryside as local policies, foreign powers, and homegrown responses to external circumstances alternately pushed the city down different developmental paths. As is suspected for many ancient Greek cities, the social community frequently extended beyond the walls of the physical city itself, incorporating rural inhabitants into the political and social apparatus of the urban community (Carter, 2006; Hansen, 2006a, p. 101–5). The composition of this community is visible through the varied architectural properties of farmhouses whose inhabitants were part of the city, providing a realistic test for the models and problems of quantitative urban studies based on archaeological data and a useful case study for understanding demography and social diversity at an urban settlement.

Farmhouses and Their Variation

Although discussions of ancient Greek cities’ social history often make use of large social categories such as rural inhabitants and urban residents, these labels mask significant variability. House architecture varied greatly at Metapontum, reflecting a shadow of the social diversity of these buildings’ inhabitants (Lanza Catti et al., 2014b). Some houses are single structures, while others are groups of buildings; some farms have towers incorporated into their design, while others do not; still others utilize an entirely or partly enclosed courtyard as a transition between rooms, while others use interior hallways and multiple exterior doors to facilitate communication. Architectural variation is not an idle difference to notice, as each of these organizational differences supports a range of social and cultural practices that facilitate the daily life of the household, and thus bear witness to the household diversity of rural inhabitants (Kent, 1990; Nevett, 2007b; Westgate, 2015). Architecture does not wholly reflect intent or constrain action, however, and diversity in household...
composition and organization will always supersede material variation (e.g., Nevett, 1999).

The integration of a courtyard, the segmentation of space into separate buildings, and the investment in towers all provide loci for the display and storage of wealth as well as opportunities for privacy and oversight to be built into the walls of a structure. Courtyards have the potential to provide privacy from outsiders for open air activities that require direct light and ventilation, but also facilitate control within the house, as rooms facing a courtyard can provide a panoramic view of tasks and movements between rooms (Nevett, 1994, 2007b; Foxhall, 2009; Westgate, 2015). Particularly in rural settings, the privacy of the family overall may be less important than the visibility of the members of the household who were perceived to be vulnerable (Nevett, 2005; Llewellyn-Jones, 2007).

Segmentation of spaces, both into multiple rooms and into different buildings may indicate an investment in infrastructure to oversee labor conducted by dependents, servants, or slaves. While cross-culturally segmentation in housing is associated with wealth, in the ancient Greek cultural context it has further been linked to the increasing presence of non-free dependents such as slaves and servants in elite houses whose menial labor needed to be segregated from the free family (Jameson, 1990; Kent, 1990; Westgate, 2015). Incorporating multiple rooms also enabled the flexible gendered use of space throughout the day and as visitors came and went (Nevett, 1994; Westgate, 2015).

Towers in houses reflect concerns about the protection of wealth and appear as a phenomenon across the Mediterranean during the fourth century BCE. They are especially common in intensive agricultural regimes where the storage of cash crops and the large-scale use of enslaved agricultural laborers went hand in hand (Morris and Papadopoulos, 2005). Regardless of whether particular rooms were used by freed or unfree people or what forms of wealth were stored in locked rooms, architectural developments at the level of individual farmhouses indicate the socioeconomic diversity present in the rural community at Metapontum and the varying strategies individual households adopted to manage and utilize their present resources (Carter, 1990; Nevett, 2005, 2007a).

Following the work of Carter, Lanza Catti, Swift, and others studying Metaponto, I gathered architectural descriptions and chronological information on 11 houses (Table 1) (Carter, 2006, p. 136–53; Lanza Catti et al., 2014b). These houses are primarily published architecturally, with the plan, types of rooms, major architectural features, and occasional finds emphasized. Thus, courtyards, towers, and the segmentation of space into buildings are the most reliable features documented for each household, and are the focus of this study. Investment in processing or storage installations, assemblages of food or textile production implements, and the arrangement of activity areas in space could also be examined for the same purpose, but were not readily available for the houses considered in this study.

The dates for these 11 houses range between the 6th and 3rd centuries BCE, with many of them broadly dated. Following standard practice for comparative urban studies, this information would normally be used synchronically by assuming that the houses present a typical view of the range of houses, and by extension the material traces of households. It is also a necessary assumption in many cases, as otherwise it is difficult to gather sufficient evidence about demography (Bagnall and Frier, 1994, p. 31–52). If we make this assumption, it is possible to classify rural houses at Metapontum according to an architectural typology that distinguishes between the presence and placement of the courtyard in each house and the presence of a tower, as these are likely socially meaningful following the discussion above. It is also possible to distinguish farmhouses by the number of distinct buildings on the same plot or sharing a courtyard. The way each house was assigned to a different category can be seen in Table 1. While this study limits itself to a synchronic approach, a diachronic approach would provide another extension to existing urban scaling studies in the Mediterranean that could be carried out in future work.

Eleven houses categorized into four architectural types or three different numbers of buildings is not a large sample to investigate, but studying empirical data at the household scale makes it impossible to throw out samples solely because they are small. It is necessary to distinguish when small samples sizes are likely to be representative of the unobserved community and when they are too small to meaningfully represent the community’s variability. In short, increasing the detail of studies of ancient Mediterranean urban studies requires consistent measures of diversity that can suggest when a small sample size captures the breadth of material culture from a community and when it is too small.

I believe this is possible by estimating the diversity of the houses classified according to a socially-meaningful typology within a community using a statistical tool developed by DeDeo et al. (2013). By assessing whether a sample closely matches the diversity of its parent population, it is possible to reject a small sample as being inadequately informative when classified in a given way, or to accept the possibility that it is representative. As with all hypothesis testing, we can only nullify the assumption that a sample is representative, but this nevertheless provides a useful check on which small datasets are good to use in community demography and which are not.

### MATERIALS AND METHODS

#### Measuring Diversity

Measuring diversity, by which I mean the total number and frequency of distinct observed types, requires a well-structured methodology. For categorical data, including the frequency of household structures or house layouts, there is no assurance that an archaeologist has observed all categories of material in a given sample for a particular way of dividing up material. To compare one assemblage with another, diversity has been dealt with by subsampling from larger distributions, a procedure which tests whether assemblages of different sizes are comparably diverse by checking how many types would show up in the largest sample if the sample were smaller (Kintigh, 1984). This makes it difficult to check whether our largest sample of a given phenomenon is representative of the unknowable parent population, however. These checks of relative diversity also do not address whether our
typology itself may be an inadequate descriptor of the material record (Cabaniss, in preparation).

To better illustrate the problem, I’ll introduce a toy example. Imagine that a group of graduate students working late in an office decide to open a variant of Scrabble in which players create words using letters printed on wooden tiles. In this particular variant, each letter appears on exactly five tiles. If a few of the tiles are missing, as often happens, someone may propose the well-intentioned but tedious idea of counting all the tiles to assess which ones are missing. While this may be feasible, it may also be impractical to check all the tiles because of constraints on time (in addition, archaeologically speaking, to constraints on funding, cultural heritage priorities, etc.). Instead, it will be easier to count and sort a sample of the tiles, as long as there exists a methodology to link this smaller sample to the total population. Arbitrarily, the graduate students may decide on a given number to count, say forty. From this largest sample it will always be possible to check whether smaller samples are consistent, but a relative methodology like Kintigh’s described above will never reveal whether the smaller tile set is consistent with a distribution containing the average size in bits (ones and zeros) required to optimally encode the output of the distribution, or alternatively how unlikely each individual category is to be observed, again in bits. For a typology with $k$ total categories, the entropy is maximized when each category is equally likely, at which point it takes the value of $\log_2 k$. This makes it a useful way to measure diversity in order to compare samples regardless of their size, as its value depends on the number of types and their relative frequency, rather than on the largest sample observed, removing the ultimate constraint in Kintigh’s framework that only lets us compare from bigger samples down to smaller samples.

$$H(p) = \sum_{i=1}^{k} -p_i \log_2 p_i$$

A further extension is to draw subsamples that are the same size as the original sample. While unintuitive, this bootstrapping procedure creates new samples of the original sample size and can thus test congruity between the original sample and estimates drawn from it. This is the EESB proposed by DeDeo et al. and implemented in the python library THOTH (DeDeo, 2013; DeDeo et al., 2013). The logic behind this procedure works as follows for the case of our Scrabble tiles (Figure 1). Having already drawn a sample of forty tiles from the Scrabble box, we construct new datasets that are similar to the original data by repeatedly sampling randomly with replacement. For each new dataset we wish to make, we draw one tile from our group of forty, record the letter on it, and return the tile back to our sample, until we have repeated this forty times. We now have a list of letters that is the same size as our sample, but on average different in key ways: letters that only appeared on one tile in our sample are very likely to be missing, while letters that appeared on multiple tiles in our sample are likely to be represented more frequently. When a letter completely drops out of this new sample, that type is no longer observed. Each time that a letter drops out and other types become more frequent, the entropy of that dataset is no longer observed. Each time that a letter drops out and other types become more frequent, the entropy of that dataset is no longer observed.
then note the decrease in entropy between our original sample of tiles and our new artificial sample that we generated randomly with replacement.

Repeating this procedure thousands of times, we can estimate the systematic decrease in entropy between the original sample and the newly generated samples. Assuming that our original sample of forty tiles was effectively selected from all the tiles in the Scrabble box in a similar way, this systematic decrease in entropy provides a good estimate of how much less diverse our sample of forty tiles was compared to the total set of tiles in the box. If this difference is large, it suggests that our real sample of tiles doesn’t capture the diversity of all the tiles in the box, since on average our artificial datasets based on our real sample don’t capture the diversity of the sample. If this difference is small, it suggests that our sample reflects the diversity of the population well, since our artificial samples reflected the diversity of our real sample.

In effect, the EESB describes how different our sample of tiles is from the population in the box in terms of the observed frequency of letters. If the two differ radically from each other, a large difference in entropy would be expected, effectively indicating a lack of representativeness based on the poor correspondence between the entropy of artificial datasets drawn from the sample and the sample itself.

For our graduate student friends, this means that they can decide to sort and count a smaller group of tiles and then calculate how much smaller the entropy is compared to what they expect, namely how much diversity, inequality, or heterogeneity is missing from their sample. By comparing this EESB estimate with the expected value of an even distribution of the alphabet, which would be $\log_2 26 \approx 4.7$, the graduate students can assess whether the population entropy estimated from their small sample is consistent with the entropy for an even distribution over the letters of the English alphabet.

For archaeologists interested in making inferences about the diversity of collections of archaeological entities such as houses, the EESB provides a way to assess how well the entropy of the observed sample represents the unobserved population. Future statistical work will focus on determining what constitutes a sensible cutoff for representativeness. As a simple heuristic, however, a sample can be rejected if the estimated population entropy is greater than the maximum entropy achievable in a given typology. The maximum entropy possible for a typology with $k$ categories is reached when each type is equally common, reaching a maximum value of $k$. If the estimate of the population entropy is less than this maximum, the small sample categorized according to that typology appears to be consistent with the overall population categorized by the same typology. If the EESB average is higher than this upper limit, however, it would suggest that there are further missing categories that the archaeologist did not observe, as the population entropy is higher than the maximum possible entropy of this typology. A high estimated population entropy implies that the sample is too sparse given the typology under consideration, and either more data must be collected or the typology must be changed to better mesh the archaeological data with the research question. This test only catches one type of error, and future statistical work is needed to better distinguish other types of under-sampled distributions from well-sampled ones (Cabaniss, in preparation).

**Experimental Demonstration**

To demonstrate the method, I generated artificial data sets that attempt to capture the effects of different sizes of typology, different structures of samples within these typologies, and...
different samples sizes. In order to generate these data sets, I created categorical distributions with the number of categories ranging between 2 and 10, making 10 different distributions for each number and thus 90 total. These were drawn from a uniform Dirichlet distribution, a common way of generating random numbers that correspond to the probability of observing different categories of a categorical variable, such as letters in the alphabet.

I then generated datasets from each of these distributions, varying the sample size between 2 and 30 and creating 10 different datasets for each sample size. This created 290 datasets for every unique distribution for each number of categories, or 26,100 datasets in total. For each dataset I calculated the naïve entropy of the sample and then applied the EESB to calculate the mean and 95% confidence interval of the entropy for the parent population. The python code is included in the Supplementary Materials.

Rejecting samples whenever the population entropy estimate is greater than the maximum entropy \( \log_2 k \) of a typology of size \( k \) does not catch all types of errors, but does predominately separate overly small samples from a mix of under-sampled and well-sampled distributions (Figure 2). As with most statistical testing, we can only reject the sample as not representative of the parent population at a given level of description—it is never possible to prove that the sample is definitively representative.

The EESB is a useful test for checking the relationship between a small dataset categorized according to some meaningful typology and whether the typology is a good descriptor of the overall material. I now turn to a hypothetical application that illustrates the problem of sparse empirical observations in archaeological urban studies.

Demography and the Three Bears

The EESB is useful because it allows us to reject a small sample and a typology when they cannot be representative of a parent population. This makes it possible to take a set of houses, assign them to types according to some useful typology, and assess whether our typology violates consistency between the sample and population if it were possible to apply the same typology to the parent population. Provided that they pass this basic test small samples can provide insight for settlements where few houses have been archaeologically documented.

For an example, let us consider the archaeological record of the fictional ancient city of Goldilockopoli. While once a large city, to archaeologists only three houses are known, known each as Casa Papa Bear, Casa Mama Bear, and Casa Baby Bear. According to a local archaeologist’s typology, each of these belongs to its own particular type of dwelling: Casa Papa Bear is a good example of a large estate for an extended family, Casa Mama Bear is a typical house for a nuclear family, and Casa Baby Bear is just right for a single individual. Based on this small sample of three houses, each in its own category, a naïve suggestion would be that one third of the households at Goldilockopoli were single resident units, one third were nuclear families, and one third were extended families.

To assess whether this sample is consistent with the unobserved population, we can run the EESB on our dataset (Table 2). As intuited for such a small and sparse sample, the results are discouraging: the EESB estimate of the parent population’s entropy is higher than \( \log_2 3 \), which suggests that our sample is not consistent with a larger, unobserved population that can also be described using only three categories. Given the small sample size, the high parent population entropy estimate implies that we need to gather more data before attempting a further demographic analysis.

The EESB provides a consistency check to assess how similar a small sample and a typology applied to it are to the inferred parent population. While future work will need to expand on how to best distinguish representative from non-representative samples, for now this can be done by comparing the estimated entropy with the maximum achievable value of a typology with \( k \) categories, namely \( \log_2 k \), to reject a typology and sample as failing to capture the probable variability of the original system, requiring either a new typology, or more data. The simulated datasets demonstrated how sample rejection works in the abstract, while the toy study from Goldilockopoli showed how this method would reject a typology and sample as not accurately representing the parent population, making a specific study of households impossible for that city without more information.

With the use of the EESB in mind, I now turn back to the real case study, the city of Metapontum in southern Italy, where I will apply the bootstrap to rural houses occupied by urban households and demonstrate that they are a sample that appears to be consistent with the broader population of houses following two different typologies. Because of this, I will then proceed with a demographic calculation specific to Metapontum between the sixth and fourth centuries BCE as well as a discussion of potential implications for the study of the society at Metapontum. On this basis, I will propose an extension to existing urban demographic methods used in the Classical Mediterranean that can aid in the comparative study of ancient urbanism.

RESULTS

Applying the EESB to farmhouses at Metapontum according to a four-category architectural typology reveals that the synchronic view of the data bears the hallmarks of being a representative sample, namely the naïve entropy is close to the bootstrapped population entropy (Table 3). In particular, there are four observed categories in the present typology, and the maximum value for the entropy of a four-category typology is \( \log_2 4 = 2 \) bits, 1 bit greater than the expected entropy from the EESB. This suggests that it is not possible to reject our typology and sample as inconsistent with their inferred parent population. In short, the population entropy suggests that the typology used could accommodate the actual diversity of the countryside while more houses to be observed. A notable caveat is that the upper end of the 95% confidence interval is too high, and more statistical work will be required to construct precise tests of what thresholds should distinguish more representative from less representative samples. At minimum, this sample of 11 houses is internally consistent with the architectural diversity present in the ancient countryside. Otherwise, if the expected population entropy from the EESB had been > 2 bits, it would indicate that our typology would need more categories to document the variability of the countryside, simultaneously indicating that the sample is too small or that the typology needs to be redefined.
From the simulated data experiment, the EESB estimated entropy minus the maximum entropy for the number of categories in that distribution is plotted against the sample size. All values > 0 would be rejected following the recommendation proposed here to reject samples when the EESB entropy is > \( \log_2 k \). Blue plusses are used to represent samples where the EESB estimate was within 10% of the actual entropy of the source distribution, while a red x is used to represent samples > 10% of the actual entropy. Samples from low-entropy populations, namely those that are diverse with only a few common types, remain below the maximum entropy even when the sample size is too low to produce accurate estimates, and thus further criteria will be required in the future to distinguish representative and non-representative samples. High-entropy populations, namely those that are diverse, and have approximately equally common types, produce erroneously high estimates when under-sampled, making it easy to distinguish them from more representative samples.

<table>
<thead>
<tr>
<th>Types</th>
<th>Houses</th>
<th>Naïve entropy</th>
<th>EESB entropy</th>
<th>95% CI EESB entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>1.585</td>
<td>2.208</td>
<td>1.584–3.170</td>
</tr>
</tbody>
</table>

The EESB entropy for the three houses at the fictional site of Goldlockspoli is significantly higher than the naïve entropy and the maximum entropy of a 3-category distribution (\( \log_3 3 = 1.585 \)), indicating that the sample is unlikely to be representative of a larger population if categorized by this typology.

<table>
<thead>
<tr>
<th>Types</th>
<th>Houses</th>
<th>Naïve entropy</th>
<th>EESB entropy</th>
<th>95% CI EESB entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>11</td>
<td>1.686</td>
<td>1.905</td>
<td>1.435–2.688</td>
</tr>
</tbody>
</table>

The EESB entropy for the four-category architectural typology applied to rural housing at Metapontum is only slightly higher than the naïve entropy and less than the maximum entropy of a 4-category distribution (\( \log_4 4 = 2 \)), which suggests that the houses are consistent with a representative sample from a larger distribution when categorized with this typology.

Similarly, the sample of documented rural farmhouses appears to be representative of the diversity present in the overall countryside when the farmhouses are categorized according to the number of buildings in each dwelling (Table 4). Here most houses consist of single structures, while only two farmhouses make use of multiple structures on the same plot. Both of these farmhouses also structure interactions through courtyards, which similarly provide the possibility of separating activities and individuals. The segmentation of space into multiple structures would be a supplemental organizational technology for households at Metapontum, rather than as essential as the use of a courtyard.

Some of the variability present in the hinterland of Metapontum is diachronic, and partly the variation in farmhouses discussed here results from the changing social history of the town. Towers were not introduced until the second half of the fourth century BCE in the region, while the first courtyard house appeared in the sixth century BCE but saw few comparable structures until the fourth century BCE as well. If these architectural features are choices to facilitate privacy, oversight, or security, some inhabitants of the countryside changed their priorities through time, increasing the infrastructure implemented to control interactions compared with houses where these behaviors were not encoded in the layout of the house, although they could have been facilitated by more perishable means. As most comparative urban studies combine multiple periods of evidence in order to create larger sample sizes, I follow the same procedure here. I note, however, that chronologies are just a temporal form of typology in terms of our state of knowledge, and this same technique can be used to assess chronological representativeness or the intersection between a typology and the periods over which it is represented.

Since the EESB estimates indicate that our sample and its typology are structured like a representative sample, the material types described in the sample can be interpreted...
as representative of the population, facilitating demographic and social reconstructions of the broader urban population. Had the sample of rural housing failed this test, it would be irresponsible to continue without either adding more houses to the sample (likely at the loss of chronological or geographic specificity) or to consider a different typology that would be less sparsely documented. Instead, it is appropriate to interpret these data because of the certainty the EESB gives and move from documenting architectural diversity to interpreting the social and household diversity these material conditions partly shape and reflect, thus giving empirical basis to the study of the community at Metapontum.

**Demography, Inequality, and the Effects of Urbanism**

Demography in the Mediterranean world has largely relied upon homogenizing assumptions that all households are equal and can be treated interchangeably. These assumptions in part stem from Bagnall and Frier’s (1994) work on the Roman census records from Egypt, yet their calculation of an average family size does not merit ignoring their variability. To the contrary, their analysis of the census data highlights the wide range of family and household sizes, their systematic variation with size of settlement, and the compounding effects of household lifecycles on calculating meaningful household statistics (Bagnall and Frier, 1994, p. 53–74).

While architectural typologies fail to capture the full range of unique characteristics that distinguish houses and their occupant households, socially-informative features such as those proposed above can inform our reconstructions of household demography. If courtyards are more frequent when issues of privacy and observation are important, it is not unreasonable to infer that larger families and more household slaves and servants were foreseen, even if never realized. If towers are a correlate of wealth, and either a direct or indirect correlate of agricultural landholding and increased numbers of enslaved persons, then their presence may indicate a still greater household population. Compared to smaller single-structure farmhouses, there is a clear spectrum here of expected populations in different types of houses when a household builds a given house. A blanket average does reconstructions of the social life of Metapontum little justice, particularly given the demographic importance of including slaves and non-kin dependents into calculations of community populations (Storey, 2001, 2002).

Given the range of population estimates in the Greco-Roman world for households, we can begin to piece together some estimates for household co-resident groups at Metapontum during the sixth through third centuries. Six of the eleven houses had neither a courtyard nor a tower. If those are indeed markers of greater wealth and particular priorities concerning privacy and security, then these other houses either encoded similar ideas in less permanent means, or did not have the interest or ability to encode those priorities in the architecture. These may then be smaller or less well-off farmers, although they are unlikely to be poor (Pettegrew, 2001; Winther-Jacobsen, 2010). Bagnall and Frier (1994, p. 67–8) calculate an average family size of slightly less than five without including any slaves or non-kin members. This is an average over a variety of household forms, ranging from single residents to large multiple family households. If it were possible to better relate the material record with particular household forms, this average could be updated. Given that these houses likely were not constructed for or by poor families, it would fit their demographic reconstructions to include at least one slave and occasionally a lodger, bringing an estimate to around 6 for the least architecturally elaborated of the houses in our sample.

The remaining five houses had some form of courtyard, which may indicate that half of households needed to oversee interactions and regulate the movements of individuals within the house. This serves both as an indication of wealth, such that this family has the ability to carry out more stringent performances of gender and status roles, and of the size of the household, namely that multiple dependents, servants, or slaves would have been present. Bagnall and Frier note that in villages, household size, wealth, and the number of different types of family members cohabiting are tightly correlated, whereas in metropoleis the correlation is not as clear, and smaller families may be preferred by wealthier households; however, this difference is more than made up for by the greater numbers of slaves and non-kin lodgers or renters such that more urban households tend to be larger (Bagnall and Frier, 1994, p. 71). Such an assumption would make the average household size for this group closer to 7.5.

Out of those houses with a courtyard, one also has a tower, which may be a marker of even more extreme wealth and potentially connected to the deployment of large agricultural workforces. Morris and Papadopoulos (2005) propose that the ancient Greek mode of agriculture would rarely have used groups of slaves larger than 10 people, and so a reasonable estimate would be half of that slave population, or 9.5 persons on average. It is possible that some of these people lived outside of the estate, but the extent to which this occurred is not presently known, and likely a good target for future work.

Our representative sample of rural housing at Metapontum provides new ways to think about ancient urban communities based on the observed properties of excavated houses. The consideration of social diversity makes it possible to ask questions about systematic relationships between status, wealth, household organization, and demography as well as their frequency at individual communities in an empirically grounded manner. If we wish to calculate an average household size, it is still possible to calculate that 6.86 people are estimated to live in each house (Table 5). Considering the relationship between individuals and households, however, emphasizes the variable experiences typical of different classes and statuses within the population.

Should we wish to think about the average size of household that a person lives in, we would note that the average person lives in a household with 7.04 people, as more people live in the larger households and most households are large. Typical lodgers, slaves, and other members of the household external to the family live in an even larger household with an average of 7.54 people per household because of the tendency for larger, wealthier households to own more slaves and have the facilities for more lodgers. Incorporating variability and social diversity into demographic reconstructions extends beyond deriving a single average across all households, and instead makes
it feasible to address more complex associations traditionally near the heart of anthropological household archaeology
(Wilk and Rathje, 1982).

The record of local housing enables a deep probing of urban social networks and society. Even small samples of housing can be analyzed to define specific demographic distributions for individual cities based on comparanda that relate architectural and social diversity. Further forms of material evidence and a greater breadth of comparative frameworks can be used to analyze these small groups of households and fruitfully investigate social diversity. With this method and small datasets like the rural houses of Metapontum in mind, I would now like to turn to some applications of the EESB to studies of ancient urbanism.

**DISCUSSION**

Empirically studying the material remains of households, even when scantily preserved, directly impacts comparative archaeological studies of urbanism. Almost all studies of archaeologically documented urbanism rely on sites like Metapontum for the majority of their case studies, where an extensive publication record in English and a decades-long research program facilitate its inclusion in English-based comparative studies. Metapontum itself frequently merits inclusion in quantitative urban demographic studies of the Classical Mediterranean because some basic statistics of the city’s grid have been made accessible. Hansen (2006b) calculated the average urban house size as 215 m² for his study of the demography of the Greek political system, while Hanson and Ortman (2017) calculated a total density of 214 people per ha based on Carter (2006) estimate of around 3,000 houses in 70 ha for their study of urban scaling. It is this last study of urban scaling I wish to focus on in considering the impact of empirical methods for the study of households.

Urban scaling, a theoretical framework that applies scaling principles from physics to the study of human social forms, has postulated mechanisms that link diverse, daily social interactions within the built environment of cities to community-level properties like productivity, crime rates, energy consumption, and infrastructure expenditures for the modern world (Bettencourt, 2013). Empirical studies and theory have identified that social outcomes such as productivity, innovation, and crime grow faster than population growth because of the canalizing effects of urban infrastructure on social interactions, while densification and reduced per-capita infrastructure enable increasing efficiencies as cities grow (Bettencourt, 2013). Most urban measurements thus grow non-linearly with respect to population size. Some, such as productivity, increase super-linearly, such that per capita economic output increases through time. Others increase sub-linearly, such that the miles of streets per capita decreases as a city’s population grows. Moreover, the framework provides the scaffolding to construct other models of social interaction and compare their outcomes at the scale of cities, testing whether community-level differences in measurable properties can emerge from differences in modes of social interaction at the scale of small groups (Bettencourt and Lobo, 2016; Cesaretti et al., 2016). Social variability must be explicitly documented, studied, and incorporated into regional studies of urbanism in order for urban scaling to have explanatory power. Unfortunately, in most regions few archaeological sites have sufficient thoroughly-investigated and well-published houses to provide even a dozen examples out of the thousands of houses that were likely built over the history of a given community.

Archaeologists have recognized the potential applicability of this theory to explaining the effects of agglomeration, and archaeological data have proved useful in testing the range of its predictions (Ortman and Coffey, 2017). Archaeological studies of scaling have encountered problems with correlated proxies, however, which confound the expected effects of scaling and introduce linear relationships between variables into our search for non-linearity (Bettencourt et al., 2013; Ortman et al., 2014, 2015) Statistically, filling in missing details about individuals and small groups with average values imposes linear models on our data. For studies of urban demography, assuming an average number of people per household for every settlement assumes that the size of household co-resident groups is invariable as a settlement grows larger, masking potential urban scaling effects.

The primary phenomenon to be explained by urban scaling is diversity, as the theory takes a staggering array of social forms, communal practices, individual identities, and

---

**TABLE 5 | Demographic calculations for Metaponto.**

<table>
<thead>
<tr>
<th>House type</th>
<th>Avg. family size</th>
<th>Avg. non-family size</th>
<th>Avg. household size</th>
<th>Number of houses</th>
<th>Expected number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>No courtyard</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>Courtyard, central</td>
<td>4.5</td>
<td>3</td>
<td>7.5</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Courtyard, off-center</td>
<td>4.5</td>
<td>3</td>
<td>7.5</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Courtyard with tower</td>
<td>4.5</td>
<td>5</td>
<td>9.5</td>
<td>1</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Average people in sample 75.5
Average household size 6.88

Following the discussion of potential average household co-resident group size, the breakdown of family and non-family members can elucidate both average household sizes as well as variation in co-resident groups. Here average family size is added to the average non-family size to derive an average household co-resident group size for each type of house, assuming an average over household forms. These can then be multiplied by the number of houses observed to calculate the expected number of people living in each type of house, which in turn reflects the average household size for those houses in this sample.
material expressions that cities play host to and identifies shared, explanatory commonalities. Studying urban scaling thus requires more empirical data about variation within and between communities, one of the original goals of household archaeology (Willey, 1982). For sites where urban housing is available, techniques assessing the representativeness of small samples such as the EESB can help further empirically ground studies of urban scaling by providing settlement-specific views of household heterogeneity, sizes of co-resident groups, and potential implications for urban social networks. Hanson and Ortman’s (2017) study of urban scaling in the Classical Mediterranean already noticed this, and they have moved the research program in a more empirically-grounded direction by incorporating information on a site-by-site level rather the application of regional rules as in previous demographic studies of the Aegean (Muggia, 1997; Hansen, 2006b, 2008). For the study of individual cities and the untangling of their tightly-correlated developmental trajectories that tie together area, population, wealth, and resource consumption, it is necessary to break down our data to the smallest empirical analytical units possible, as homogeneity confounds our object of study. Hanson and Ortman (2017) have already proposed and implemented one means of doing this, and I hope the present work will be an extension to assess how we can use small datasets of houses and other sparsely excavated material entities to expand our knowledge of ancient cities.

Empirical techniques for the study of ancient cities can be nested and extended to make the most of what limited datasets exist. Assumptions at this stage are then made more explicit and are framed at a level closer to the lived social diversity of the cities we wish to study. Step by step, our models move from ideal to real ideal, wherever possible exploiting small samples and scattered material to bring out the true richness of ancient urbanism.

AUTHOR CONTRIBUTIONS

AC designed, implemented, interpreted, and wrote the study.

FUNDING

This paper was published thanks to OPEN-AIRE funding granted to Francesca Fulminante for the Marie Skłodowska IEF Past-People-Nets 628818 conducted at the University of Roma Tre (2014–2016). This research was partially supported by the University of Michigan.

ACKNOWLEDGMENTS

Part of this paper was first written for Lisa Nevett’s graduate seminar on domestic space in the Greco-Roman world, and was subsequently presented at the Computer Applications and Quantitative Methods in Archaeology (CAA) international conference in Atlanta, GA, in 2017. I would like to thank the other students in that class as well as the audience at CAA for their discussion and input into this method. Thanks to Lisa Nevett and Sheira Cohen for reading over drafts of this work. Thanks also to Francesca Fulminante for her essential support. I would also like to thank the three reviewers whose recommendations have substantially improved the clarity and content of this paper. Any errors that remain are exclusively my own.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdigh.2019.00011/full#supplementary-material

REFERENCES


Frontiers in Digital Humanities | www.frontiersin.org 74 June 2019 | Volume 6 | Article 11
Modalità Inseidativa e Strutture Agrarie nell'Italia Meridionale in Età Romana (Bari: Editori Braidati), 197–211.


Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Cabaniss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Trajectories to Low-Density Settlements Past and Present: Paradox and Outcomes

Roland Fletcher*

Department of Archaeology, University of Sydney, Sydney, NSW, Australia

The conventional history of urban growth defines agrarian-based cities prior to the nineteenth century CE as densely inhabited and commonly bounded by defenses such as walls. By contrast industrial-based cities are viewed as more spread out and without marked boundaries. Since the 1960s a trajectory toward extensive, low-density urbanism with sprawling, scattered suburbs surrounding a denser core has been formally recognized and given various names such as megalopolis in the West and desakota in southern and eastern Asia. These sprawling industrial cities have been regarded as a unique derivative of modern phenomena such as mechanized transport and the commercial property market. However, this set of premises are not valid. The agrarian-based world also contained dispersed, low-density urbanism—on its grandest scale, the vast circa 1,000 sq km urban complex of Greater Angkor and the famous Maya cities of lowland Central America with maximum areas of about 200 sq km. The Maya only used pedestrian and riverine transport so the conventional transport explanation for industrial dispersed urbanism is at best partial. There was another trajectory to extensive, low-density settlement forms for places which were generally <15–20 sq km in extent but could on rare occasions reach areas as large as 40 to 90 sq km. Famous examples are Great Zimbabwe, Chaco Canyon and the European oppida of the late 1st millennium BCE. No-formally agreed term is available to refer to them. I will refer to them by default as “Giants.” The three trajectories to low-density settlement form redefine the history of settlement growth and the meanings of the term “urban.” Worryingly, none of the successive low-density settlements derive from any of the low-density cases of the preceding trajectory. Neither Angkor nor the Classic Maya cities have any connection to the industrial low-density cities. By contrast compact cities, the epitome of the obsolete definition of cities display continuity to succeeding urban forms over several 1,000 years. The implications for modern, giant, low-density cities are ominous.

Keywords: low-density, urban, past, present, future

INTRODUCTION

Since the 1960s a trajectory toward extensive, low-density industrial urbanism with sprawling, scattered suburbs, surrounding a denser core has been formally recognized and given various names such as megalopolis in the West and desakota in southern and eastern Asia (Gottman, 1961; McGee, 1967, 1991; Doxiadès, 1968; Angel et al., 2005, 2012; Morrill, 2006; Angel, 2012).
These sprawling industrial cities (Figure 1) have been regarded as a unique derivative of modern socio-economic phenomena, such as mechanized transport and the commercial property market. The conventional history of urban growth defines agrarian-based cities prior to the nineteenth century CE as densely inhabited and commonly bounded in some way, for example by topographic features or walls. By contrast, large industrial-based cities are viewed as more spread out and lacking marked artificial boundaries.

However, this set of premises is not valid. The agrarian-based world also contained dispersed, low-density urbanism. On the grandest scale these include the vast urban complex of Greater Angkor (Figure 2; Groslier, 1979; Pottier, 1999; Fletcher et al., 2003; Evans et al., 2007), which at its peak in the twelfth century covered ~1,000 sq km, Anuradhapura and Pollonnaruwa in Sri Lanka which ended between the eleventh to thirteenth century CE (Devendra, 1959; Gunawardana, 1971; Coningham and Gunawardhana, 2013) and the famous Classic Maya cities of lowland Central America (Figure 3; Sharer and Traxler, 2006), with maximum areas between 100 and 200 sq km, which faded away in the ninth century CE. The Maya only used pedestrian and riverine transport so the conventional transport innovation explanation for dispersed urbanism in industrial societies is at best partial. The presence of sprawling suburbs in the agrarian cities suggests that the old models of urbanism are both incomplete and potentially problematic. The existence of the agrarian-based, low-density cities also specifies that the low-density pattern is a usual feature of human behavior rather than a unique and anomalous aspect of “strange” industrial urbanism. The extension of that understanding is that there should be other such settlements of varying functions and magnitudes.

Since we know that some hunter-gatherer communities live in extensive dispersed camps (Fletcher 1), as for example, among the Australian Aboriginal communities of the deserts of Australia (Figure 4; O’Connell, 1977; Whitelaw, 1991) and by the Ainu of Hokkaido (Watanabe, 1973), this must be an essentially universal human behavioral characteristic. Conventional agrarian “village” communities, likewise use this settlement form. Large regions of West Africa, inhabited by the Kofyar (McNetting, 1968), the Tallensi (Figure 5; Fortes, 1945) and the Lobi-Dagati societies (Figure 6; Goody, 1956) among others, are covered by vast dispersed villages where the houses are widely separated and each is surrounded by fields1. A “village,” occupied by one self-defined community, can extend across several kilometers. This form of settlement was also very well-known in some regions of Europe and has featured in settlement pattern geography (Cerne, 2004) at least since the mid twentieth century, for example in Hoskins famous “Making of the English Landscape” (1955).

We therefore know of low-density, dispersed settlement forms ranging from several 100,000 sq km, such as the megalopolis, to dispersed hunter-gatherer camps covering several hectares. Just as we have spectacular, giant industrial cities in the range up to more than 100,000 sq km and vast, agrarian urban, low-density cities, covering as much as 100 to a 1,000 sq km we should therefore also have a set of extensive dispersed agrarian-based settlements at the next smallest order of magnitude. We already know that dispersed agrarian “villages” exist with areas larger than a square kilometer. So we should also expect that this magnitude of dispersed settlement would also have produced settlements of proportionately spectacular size and form covering from several sq km to tens of sq km. Famous examples of this third size-range of spread-out settlements (Figure 7) are Great Zimbabwe (Garlake, 1970; Pikirayi and Chiriikure, 2011; Chiriikure et al., 2016, 2017), Cahokia (Fowler, 1989; Kelly and Brown, 2014; Baires, 2015; Pauketat et al., 2015) and the European oppida of the late 1st millennium BCE (Collis, 1984; Sievers and Schönfelder, 2012; Moore et al., 2013; Poux, 2014). It turns out that some very rare example even range up to 40–70 sq km and perhaps larger (Figure 8) such as Gelonas—a Scythian gorodische (Shramko, 1987, 2012; Murzin et al., 1999; Zöllner et al., 2008) and the vast area of Chaco Canyon (Vivian, 1970; Wills and Dorshow, 2012; Drake et al., 2014; Price et al., 2017) in the SW of North America. As yet no-formally agreed term is available to refer to this class of settlement though some are contentiously called urban or else rejected as urban, either periodically or at the same time. As Taylor remarked of the Tripillya sites (1987, p. 4) are they either “cities” or “just overgrown villages? (see Zbenovich, 1996; Nebbia, 2017)” These kinds of settlements are plainly not “just” villages (Chapman et al., 2014), as a set of dispersed, agrarian villages is already known to exist. These are something more. I have previously called such sites, often with massive constructions or a distinctly ordered spatial layouts, “No-Name” settlements2. More recently they have been referred to as Anomalous Giants. Since they are only anomalous relative to our current restricted “camp-village-urban” taxonomy of settlement form, and not due to some inherent anomalous strangeness of their own, a simple way of referring to them is needed. “Giant” sites, as proposed by Chapman and Gaydarksa, may serve as a default for now, because the term avoids a definitive ascription of their role or categorization.

These Giants have not been “seen” as a distinct set of settlements because dispersed occupation settlements in general were not “seen.” This is not a case of data not being known. Rather it is that the compact village has been a very strong normative model of “the village,” despite dispersed villages being known in academic literature since at least the early-mid twentieth century. In addition, the Classic Maya urban dispersed pattern had been seen in the Ricketsons’ surveys around Uaxactun in the 1930s (see Black, 1990), was understood in the 1960s once Willey’s program of research developed (Willey, 1956, 1965; Sabloff and Fash, 2007) and was formally described by Puleston (1973) for Tikal in the 1970s. This information was available to be combined with the observation of industrial dispersed urbanism to create a new view of what was feasible for urban settlements. But facts do not

---

1 As I well know from living in such settlements in the early 1970s, when I was doing village household surveys in Ghana for my PhD research.

suffice, in themselves, for the phenomenon they represent to be perceived and understood. The data need to be seen in a larger context which gives them meaning. This can be both a theoretical context and a particularly vivid empirical context, though the empirical context is likely to come about because of a conceptual proposition which directs attention to collecting that contextual data. The form of the “anomalous giants” was not a topic of focus because low-density settlements were not a topic of discussion and the giant agrarian cities which have since directed attention to that topic could not readily be recognized.

**WHY WERE LOW-DENSITY DISPERSED AGRARIAN-BASED CITIES NOT OBSERVED?**

Compact urbanism was seen as the only form of agrarian-based urbanism by scholars trained in Europe. Their understanding of the Islamic world, Mesopotamia, India, and China strengthened this view. The viewpoint was good common-sense, because European scholars simply did not experience, and nor could they readily “see” any dispersed urban settlements which had predated the later nineteenth and early twentieth century. It was
FIGURE 3 | Maya settlement of Sayil, Mexico, Terminal Classic, tenth century CE. Permission to published granted by the copyright holder, the Middle American Research Institute.

simply self-evident that cities were compact because all known agrarian-based European cities were compact and were delimited in various ways, as were most other cities of the past in the Old World, outside Africa, and also in the Mesoamerican highlands and along the west coast of South America. For the social and intellectual traditions of both India and China the compact, bounded city was normal.

Industrial urbanism could be considered to be different—more spread out—and with a ready explanation of their form in the mechanics of transport, a new economy and the role of the real estate market. Therefore, because cities were “traditionally” compact, modern sprawl was unique and idiosyncratic, and its form was therefore uninformative about the past, and did not lead to the question—were there dispersed low-density urban settlements in the agrarian world? Plainly, industrial cities have to be called urban and are sprawling. They, therefore, could in principle have specified that if they were “urban” then urbanism in general would include dispersed and patchy forms of occupation. But the empirical evidence prior to the 1960s also did not indicate that there was any topic to which that proposition could attach, though Pagan in Myanmar (Figure 9; Pichard, 1995; Hudson, 2004) was a potential exception.

Most of the large, low-density agrarian cities which we know from archaeology are in primary and secondary tropical forest. In consequence their layout is hard to see. And due to the use of domestic housing built mainly from organics, the residential landscape is not readily apparent in most tropical regions with the exception of the Maya sites where the lightly built houses were raised on stone rubble platforms. The remains of a dispersed urban settlement were therefore generally less likely to be observed by walking through it. In addition, the majority of agricultural, literate, low-density cities had ceased to function and be fully inhabited by the mid-2nd millennium CE before Europeans could see them. And even when a foreign observer saw one—Zhou Daguan being the notable example in 1295-6 in Angkor—he “saw” it, as his own Chinese cultural familiarity would lead him do, as the walled “city” of Angkor Thom, with some huge structures such as Angkor Wat well-beyond the walls (Daguan, 2007). Likewise, for European scholars arriving in the largely abandoned landscape of Angkor in the later nineteenth century, they also saw Angkor Thom and the other great temple enclosures as the equivalent of European medieval towns because they were of similar areal extent (Figure 10).

The rare case of Pagan is highly informative about classificatory pre-designation. The clearly visible 90 sq km expanse of dispersed monuments (see Figure 9) which is in an open, dry landscape was viewed simply as a small walled city because of the palatial enclosure near the river, with temples scattered around it. Classificatory expectation had precedence over what could be seen. Likewise, from the 1930s to 60s, the Maya cities were seen as temple clusters isolated in the jungle (Webster, 2007). And when the vast dispersed extent of these cities was empirically recognized the explanatory priority was to legitimate them as relatively “dense” areas of occupation in order to validate their urban status. Only since the spread—out patterns was clearly recognized and had been studied in detail did the focus shift decisively to their open extent (e.g., Sabloff, 1990; Graham, 1999; Isendahl, 2012). Yet until recently they were still regarded as so unusual, globally, that a case to redefine agrarian “urbanism” to include dispersed, low-density form would have had a hard ride. Likewise, at Angkor Groslier had identified the spread-out distribution of occupation and shrines in Angkor by the 1950s and 60s. He published his model of extended suburbs linked together by road embankments and canals in 1979 (Figure 11) as part of a study of the areal expansion of Angkor from the ninth to the fourteenth century. However, this insight did not gain traction either in Angkorian studies or with SE Asian regional specialists. The walled enclosure of Angkor Thom continued to be perceived as the “city” and the prior assumption (Briggs, 1951 for example) that Angkor had just been a succession of small, walled “cities” from at least the ninth century—and also the Goloubew interpretation of CP807 (see Pottier, 2000)—was maintained. No change could occur in interpretation from the 1970s into the 90s because field archaeology was impossible in Cambodia due to severe political crises. New empirical data could not be obtained and foreign archaeologists could not assess reinterpretations of the urban landscape. Only in the 1990s did Christophe Pottier start his PhD research on the landscape of Angkor, following up on Groslier’s insight and mapped the southern suburbs of Angkor in detail (1999). He also showed that the mid-nineth century capital of Hariharalaya (modern Rolous) did not have a boundary wall and that the major temple enclosure around the Bakong did not encompass much occupation (Pottier, 2005). Furthermore, he showed that Goloubew’s proposition that CP 807 was one corner of a walled city dated to the late ninth century was incorrect...
WHAT DIRECTED ATTENTION TO LOW-DENSITY URBANISM?

Pottier’s maps, based on the remote sensing images of the FINMAP aerial photographs and his own ground surveys, combined with the validation of Groslier’s intellectual status provided the first decisive images of an Asian agrarian-based, dispersed urban complex. His precise arguments about the form of Hariharalaya confronted Angkorian specialists with a debate about urban form, related to a historical process and to the established specification that Hariharalaya was the “model” for Angkor and therefore needed to be addressed. My own arrival in Angkor because of very specific research issue raised by my own theoretical work (see below p. 7–9) triggered collaboration with Christophe and a follow up on the initial space borne and aerial radar surveys of 1994 and 1996 by NASA—JPL. What is significant is that neither of those initial remote sensing radar surveys, in themselves, led the scholarly world to recognize the urban form of Angkor, despite Groslier’s work and the paper by Jacques (1978) where he shows an extended North but then does not produce later maps with it marked.
The 1994 survey from the Space Shuttle—courtesy of WMF and NASA (Figure 12) was, however, the empirical trigger for me to conclude that Angkor was indeed a giant low-density city. Up to that time I too had perceived Groslier as somehow describing the landscape of a region. The following, 1996 AIRSAR radar survey did not trigger a re-perception of the form and extent of Angkor because it was a partial strip across the southern half of Angkor and was used by its instigator, Moore (2000), in an erroneous claim that radar was primarily a useful tool for finding undiscovered temples. As a consequence, her focus emphasized isolated specific structures in a landscape, not the relationships which could be clearly seen between sectors of the occupation, the infrastructure and the rice fields of Angkor, as Pottier had emphasized in 1999.

The basic problem was that the sheer scale of Angkor militated against seeing it as whole and made studying it exceptionally difficult, especially on the ground in areas of either dense ground cover or extensive rice fields, in a region which was extremely wet for half of the year. Aerial photo coverage could give real assistance for surveys but did not in itself offer a vivid visual image of the whole because the coverage was, as was normal, divided into tiles of varying clarity and contrast. The 1994 space radar image did, however, spectacularly cover the whole of the Angkor urban complex, for the first time in one image from the Kulen hills to the lake and from the Puok river to the Damdek canal. The Great North Canal is starkly visible, showing for the first time the full extent of the infrastructure connections to central Angkor3. What followed from the collaboration between the University of Sydney and the EFEO was an international program to carry out a 7,000 sq km GEOSAR survey over Angkor to create the map of the entire area of the Angkor urban complex (see Evans, 2007 and Evans et al., 2007). That presented, in very consistent, recognizable detail and beautiful color the extent, configuration, structure, and economic food source foundations of the entire urban complex of Angkor. Greater Angkor could then be designated (Fletcher et al., 2003), avoiding terminological conflict with those scholars who are primarily and properly focussed on the central area of Angkor—“Angkor” in the usual parlance. A gratifying collaboration developed between Damian Evans and Christophe Pottier to create the digital map of Greater Angkor, its water catchment and the wider local region, which was published in 2007 (Pottier and Evans, 2010). A new edition has recently been completed by Evans team in the EFEO.

Once the lidar coverage of much of the large Classic Maya city of Caracol was completed in 2010 (Chase et al., 2010, 2011, 2012; Figure 13), the actuality of dispersed, low-density agrarian-based urban landscapes, worldwide, was unavoidable. The lidar image corroborated the Chases’s ground surveys, carried out of over many years, which had recorded extensive terracing and widely distributed domestic units scattered between roadways that linked the central area to the peripheral nodes of the urban complex. The value of remote sensing using lidar and its key role in studying dispersed urbanism in tropical forests was further reinforced by the 2012 lidar survey of the 2–300 sq km area of the Heritage park in the middle of Greater Angkor by the KALC project (Evans et al., 2013). As well as demonstrating the value of lidar for revealing detailed urban landscapes under dense tropical forest it also showed how profoundly the central area of Angkor had been modified between the ninth and twelfth century, decisively showing that the twelfth century grid was present outside Angkor Thom as well as within its late twelfth century walls (Figure 14). The 9 sq km of Angkor Thom could not, therefore, be the city of Angkor, as such, since the rectilinear route grid and small pond configuration extended over more than twice that area outside the walls. Furthermore, the small pond residential space overlaps with the distribution of the much larger ponds which are the dominant pattern of the suburban residential landscape of the whole of Greater Angkor. A residential continuum existed across which hundreds of small shrines and many thousands of occupation mounds and ponds.

3I remain deeply grateful to the staff of the National Air and Space Museum in the Smithsonian Institution for saying to me after a seminar, “by the way, have you seen the radar image of Angkor from the Space Shuttle?”
FIGURE 7 | Giant sites compared—Old World. Prepared by Kirrly White.
FIGURE 8 | Giant sites compared—New World. Prepared by Kirrily White.
FIGURE 9 | Pagan, Myanmar eleventh–thirteenth century CE. Permission to publish granted by the creator and copyright holder, Bob Hudson.

FIGURE 10 | Angkor Wat, Angkor, Cambodia twelfth century CE, compared to sixteenth century Leiden, Netherlands, Europe. Prepared for the Greater Angkor Project by Martin King and Kimly White.
were distributed both within the central area and out across the landscape of the urban complex.

**LOGICAL CONSTRAINTS, IN THE TWENTIETH CENTURY, ON COMPARISON ACROSS DIFFERENT MAGNITUDES OF SETTLEMENT SIZE**

The non-recognition, until the start of the twenty first century, of low-density dispersed settlements in archaeology, as a systemic, global phenomenon across many scales of settlement size, is curious since they were well-known in anthropology and landscape studies by the mid-twentieth century, and the Classic Maya cases were clearly visible by the 1960s, as was the identification of the class of industrial urban “sprawl” i.e., dispersed, patchy urbanism in the modern world, as megalopolis, by Gottman in 1961. The term had even been used by Geddes in 1915 and by Lewis Mumford in his 1938 book “The Culture of Cities,” in relation to the vast, industrial urban conurbations, such as the Ruhr, which Mumford described as the first stage in urban overdevelopment and social decline. The question could therefore have been systematically asked, much earlier than it was, about whether there had been previous, dispersed urban settlements. Logically it should have been, since dispersed, low-density villages were well-known in Europe and appeared to have been a long-established settlement pattern. So, in principle, if agrarian villages and industrial cities could both take this form, variants for other socio-economic systems with the same spatial configuration could have been envisaged. A key deterrent was presumably, Childe’s enormous prestige, and his “definition” of urbanism which included dense occupation, even though Childe also remarked on the difficulty to defining urbanism (1950), and the negative connotation applied to the initial expanding industrial conurbations.

As is apparent, empirical problems of observation do not suffice to explain the lack of an integrated recognition of the dispersed settlement form. The general phenomenon of dispersed occupation was known and the conclusion could have been reached that such a settlement form exists and has existed, long before the beginning of the twenty first century when remote sensing made it readily and spectacularly apparent. However, in addition to the combination of specific assumptions about the nature of urbanism and the practical issues of observation outlined above, a suite of broader conceptual premises was also in use up to the later 1970s which militated against the necessary cross-comparisons. In order to articulate a general proposition about human residential behavior being distributed across a wide range of occupation densities and to facilitate cross-comparison between settlement types of different spatial magnitudes, some basic assumptions about the relationship between classes of settlements were required as well as changes in fundamental assumptions about the relationship between materiality and sociality.

Only when stage theory was no longer tacitly accepted as self-evident and when materiality and sociality began to be decoupled after the 1970s could the segregation of the
characteristics of the industrial cities from those of agrarian cities no longer be logically specified. This is a structural problem. Even in 2006, Bruegmann’s book “Sprawl: a compact history” triggered no sustained engagement either for modern urbanists with the past or for archaeologists with sprawling industrial urbanism. The fundamental issue was that a generalized, linear evolution was the dominant model of culture until the 1970s despite the recognition of multi-linearity by Steward and others decades earlier. Stage Theory has been, and is still, a convenient way of managing the diversity and the vastly differing magnitudes of cultural behavior. But it logically emphasizes the significance of and the interest in difference between stages and presumes strong correlations between sociality and materiality within each stage, largely precluding cross—comparison between settlements of differing magnitudes and social function.
The assumption of direct material—social correlations was also problematic because it marginalized the material as an epiphenomenon of sociality. Different material expression were presumed for different social conditions so the generalized similarity of dispersed occupation with many different kinds of sociality was not a focus of interest. In addition, large patterns of materiality, such as settlement form, were not a concern of social anthropology for which the settlement was predominantly a backdrop for the active expression of social meaning. With a strong focus on sociality, the concepts and information that were transferred to archaeology did not emphasize general settlement form but primarily directed attention to specific house forms as the corollary of particular social systems. As a consequence, the knowledge of dispersed villages in social anthropology did not generally percolate across into the teaching programs of English-speaking archaeology in the middle of the twentieth century.

The pragmatic needs of archaeological research and analysis seem to have diverted attention from the information readily available in the study of land use in Europe, and the precise observations by archaeologists in Europe that in some settlements the buildings were placed far apart e.g., Danubian sites (Soudský, 1962) and that in others they were very close together as in Skara Brae (Childe, 1931) and the Iron Age lake villages of Europe (e.g., Clarke, 1972) or in Catal Huyuk (Mellaart, 1967; Hodder, 2006). The focus in archaeology was primarily on compact settlements, despite the experience and expertise of the Mayanists, suggesting both that the compact settlement had become a normative type and that the compact settlement was a meaningful, practical, and theoretically substantial topic of inquiry. For the pragmatics of archaeology a compact site is an obvious and well-defined

---

4My own knowledge of these kinds of settlements was due to the happenstance that I met Meyer Fortes, and was taught by Jack Goody, in Cambridge who both assisted me to work in the Tallensi and LoDagaba regions of Ghana on the village surveys which were part of my PhD research.
entity to study. In addition, the work of the landscape economy theorists such as von Thunen (see Chisholm, 1972) offered access to understanding the distribution of crop production, estimating food yields and thence calculating populations. And because archaeologists wanted to find ways to estimate the population size of past communities a consistent settlement area—population correlation was being sought from the 1930s to the 1980s (see commentary in Hassan, 1981). Essentially, normative, presumably modal densities were envisaged—such as an average of 100 p/ha for agricultural urbanism. Several critical papers on the issues of population estimates e.g., by Postgate for Mesopotamia (1994) made that problematic. To do population estimates for dispersed settlements requires some combination of detailed economic data such as a rice field distributions and visible indicators of domestic water supply and the temporal structure of the settlement (Fletcher et al.,

![Figure 15](image1.png)


![Figure 16](image2.png)

2003, p. 116–117, Hawken, 2011; Klassen, 2018). In the Maya settlements house mounds are generally visible. A sophisticated archaeology of domestic populations developed. The recognition of the problem in Maya sites that some households were not very visible on the surface (Johnson, 2004) illustrates the intellectual vigor and critical articulation of inquiry in a region with dispersed occupation.

The wide range of occupation densities in human communities world-wide did not resonate in the discipline—though its reporting (e.g., Fletcher, 1981, 1995) triggered
no particular opposition. The information had little or no academic traction. The paradox is that archaeology has long recognized differential density of built space, has reported it for archaeological sites and has known of or had access to examples of such behavior for many decades. The key implication is that the dominant explanatory logic and the practical purpose of getting on with doing archaeology, demonstrates again that facts do not in themselves provide insights that lead to theories. Instead, pragmatically, such facts are simply anomalies to be left aside until some

![Graph 1: Types of trajectories on the Interaction-Communication matrix. Prepared by Kirrily White.](image1)

way to manage and integrate them can be envisaged (see Medawar, 1967).

**INTERACTION-COMMUNICATION MATRIX**

What we have tended to focus on are the ways in which interactions are managed socially and what kinds of verbalizable meaning are communicated by the messages transmitted within a community. However, when the material is regarded as an operator in its own right with consequential effects, because of its own characteristic, rather than as a consequence of or a carrier for the sociality with which it was associated, logical cross-associations can be made between settlements of very different magnitude and form (Fletcher, 1995). A basic example is that materiality manages interaction by providing barriers of various sorts and by assisting the transmission and retention of signals (Fletcher, 1995, p. 126–151). When therefore, instead of sociality and verbalized meaning, the degrees of intensity of interaction and degrees of adequacy of communication are considered as factors in their own right in community behavior, global cross comparability is feasible, at a very basic level\(^5\). The key is to view interaction as a condition which varies with occupation density, becoming increasingly more stressful the higher the density, and to view communication as an activity which becomes less effective with increasing distance, for any given means of communication. Settlement extent and occupation density are therefore consequential. An Interaction-Communication matrix (Figure 15) of density (Density—P/A) plotted in relation to community size (P) therefore allows, in principle, all settlements to be plotted relative to each other. When large numbers of settlements are plotted both globally and from national census data back to the nineteenth century (e.g., the Indian and Japanese national census data) it becomes apparent that human communities operate across an extremely wide range of residential densities. Even in densely inhabited countries like India and Japan the majority of towns have densities below 30–50 p/ha (Figure 16). And hunter- gatherer settlements range from residential densities of over 1,000 p/ha (e.g., the !Kung and the Efe) to <5 p/ha in some spread out—Ainu winter settlement camps—(Fletcher, 1995, p. 80; Whitelaw, 1991).

Interaction limits: different upper operational densities are apparent for recent mobile and known sedentary communities (Figure 17). At lower densities the areal extent of settlements can be vastly larger than the equivalent for the largest dense settlements of the same kind of socio-economic system (see Figure 19). This is the phenomenon which the IC matrix brought into focus, which directed my attention to Angkor and led to the proposition that it was a low-density city. This characteristic of low-density settlements is a crucial feature of cultural processes because it allows them to exceed the maximum operable extent of a compact settlement area for a given communication system (Fletcher, 1995, p. 117, 121, 124).

Communication limits (Figure 17): the biggest compact literate, agrarian-based cities are in the 70–100 sq km range with populations around 1 to 1.5 million (Fletcher, 1995: p. 84–87, 130). Though this area extent was attained by Chang’an in the eighth century CE (Xiong, 2000) it was not exceeded globally until the first half of the nineteenth century by London (Ackroyd, 2000). A Communication limit is therefore very constraining. While Angkor and Caracol, among others, could drop to low density and exceed the maximum area allowed by a given communication system for compact settlements they could not shift back up to a higher density form and could not apparently maintain either themselves or their extended settlements networks. The research at Angkor was initiated to investigate these issues and has provided reappraisals of its history (Fletcher, 2012; Lucero et al., 2015; Penny et al., 2018).

Over the past 15,000 years there have been three Communication limits identifiable by long period of stasis in the largest compact settlement sizes and by characteristic frequency distributions of the areal extent of settlements (Fletcher, 1986). The maximum settlement size ranges for compact settlements which have prevailed over several millennia, are about 0.7–1.0 ha for mobile communities, then about 70–100 ha for agrarian settlements and then 70–100 sq km for the great compact agrarian imperial capitals. Behind these successive C-Limits the usual distribution of the areal extent of site/settlement area is, respectively, below 3,000 sqm (relative to the 1 ha limit), 30 ha relative to the 100 ha limit and 30 sq km behind the 100 sq km limit. Most site/settlements areas are in the lower third of the size range behind a C-limit (Figure 18). The industrial urban size distribution is now mainly below 3,000 sq km.

Trajectories: because the I-C matrix is a field model it can be used to display and analyze trajectories over time and the relationship between factors such as social conditions, economic circumstances and environmental processes, and the sizes of the settlements being affected by these factors (Figure 19). Most striking is that mobile communities, trend toward lower occupation densities as settlement area increases (Fletcher, 1990, 1991, 1995; Whitelaw, 1991, p. 76–80; 166; Fletcher, 1998) while known sedentary communities display the opposite trend (Figure 20; Fletcher, 1995, p. 170–177). This is crucial because it means that no one trend of density and settlement size is applicable to all human communities. The overall energetics of the trend in mobile and low-density settlements toward decreasing density with increasing community size are profoundly different from those in compact settlements where the density increases as the community size increases. In the latter communities greater stresses are placed on interaction and communication load while concurrently more interconnectedness is facilitated throughout the community. The key issue is whether central area in low-density settlements conform to the overall settlement density decrease with increasing community size (population) and settlement area or alternatively display their own trend to increased density as central area increase in population and area extent—**independent** of the overall trend to density decrease of the whole expanding settlement (see below Issues p. 10–11).

\(^5\)Traffic analysis is required, as defined in cryptography—the interception and examination of the physical characteristics of messages such as frequency, rate and internal structure to deduce information from the patterns of transmission, even if the content of the messages cannot be decoded (Callimahos, 1989).
This is consequential, because as well as density trends there are also several distinct trajectories of settlement growth and changes in overall settlement density over time. There are two, high density trajectories—one to stasis behind a C-limit and the other—which has been rare—a transition trajectory at high density across a C-limit—as for example by London in the early phase of the Industrial Revolution. A middle density band of trajectories is very varied, tends not to involve large changes in settlement size and given their constrained sizes is generally the predominant pattern for most settlements in the lower third of the size range behind a C-limit. The third trajectory (Fletcher, 1995, p. 117, 121, 124) is very distinct and drops to low-density as settlement area increases enormously. Vast expansion occurs but this is a terminal path which ends in the demise of the settlement. This trajectory, as noted above, cannot revert to higher density at the same large areal extent because the available communication system could not support coherent communication at higher density across that areal extent.

OUTCOMES

What is striking is that the same forms of socio-political organization, such as the state, operating in settlements on markedly different occupation density trajectories, lead to very different outcomes (Fletcher, 2010). While states do break down on both the higher density and low-density trajectories their longer-term histories are very different. Compact urban networks have repeatedly demonstrated enormous resilience of despite severe political changes such as the end of the Roman Empire. Specific settlements may fail but the overall network continues and urban recovery occurs within the existing system. The indications are that communities in compact settlements are extremely robust and able to make new socio-political adjustments in situ. There is long continuity in the traditions of compact settlements. Industrial London derives from Roman-Saxon London, Baghdad and Damascus have endured over centuries despite crisis and disaster. The development of early compact urbanism in lowland Mesopotamia and highland Mesoamerica has a direct ancestry in the small villages of those regions.

By contrast, the giant dispersed agrarian urban settlements of the Old and the New World had reverted to village farmland between the ninth and the fifteenth centuries and their urban networks were abandoned. The urban development of the future in lowland Mesoamerica, Sri Lanka and the eastern mainland of SE Asia formed up on the periphery of the former urban heartlands (Lucero et al., 2015). States survived in Southern Asia (Fletcher et al., 2017) but the old urban networks did not. Even more serious, there is no continuity in the development of dispersed settlement patterns. The industrial megalopolis and the desakota do not derive from or have any ancestry in the old giant, agrarian low-density cities. Nor do Angkor and Anuradhapura have any ancestry in the “Giants,” the smaller form of dispersed settlement in the 1–100 sq km range. And most of the “Giants” elsewhere in the world had no similar descendant (Fletcher and White, 2018). As ever, lowland Mesoamerica becomes critically important for our understanding of continuity in dispersed settlement systems because the Pre-Classic to Classic to early Post Classic does involve a succession of new low-density settlements, moving their main locus further and further north along the Yucatan peninsula. Only after the twelveth century does the shift to compact settlements occur around the periphery of the old urban heartlands.

ISSUES

Once the dispersed settlement patterns are recognized, the analysis of settlement behavior can shift from classes of settlement to trajectories with quantifiable magnitudes which can be related to patterns of outcome. A significant issue for the analysis of low-density urbanism is the patent accumulation of wealth, resources and innovation in the central, more densely occupied areas of these cities as population size increases. The implication may therefore be, that the model of Bettencourt et al. (2007) should apply to the denser areas of any urban settlement, worldwide across the spectrum from compact and dispersed form at all times. In agrarian—based, low-density settlements the denser central areas are simply smaller areas than can be attained by the maximum extent of higher density, compact settlements that use the same suite of communication systems. This proposition leads to an additional socio-political implication, of some consequence for the histories of dispersed urban settlements because the center—periphery “differential” in dispersed agrarian cities with very low, peripheral densities should be greater than in a compact city. This is not just that the rich get richer—it is that the whole social fabric of the denser more central areas gets “richer.” In a denser city that differential will be less and its “pay-offs” will be more accessible because the periphery is physically closer. By contrast, in a dispersed urban complex like Greater Angkor, the social differentiation between the denser center and the periphery should rise proportionately much faster for the central populace than for the peripheral populace who are also much further away from that cumulation.

What should follow is that decoupling of wealth and social integration may be more severe in dispersed cities. If this can be tested in the old agrarian low-density cities it has some implications for the consequences of increasing social wealth differentials in our urban present and in the future of our gigantic, dispersed urban agglomerations.

As well as redefining central area sociopolitical processes the form of dispersed settlements also involves a reappraisal of the settlement’s resource hinterland. In a dispersed agricultural village the extent of the occupation area is largely coincident with the community’s staple resource hinterland because the fields are around and between the houses. Likewise, in Greater Angkor there were rice-fields all the way in to the front door of Angkor Wat, creating the “Metropolis of Rice-fields” (Hawken, 2011). Caracol’s urban area, tied together by roads, secondary centers and outlier shrines is almost entirely terraced (Chase et al., 2011). Many years ago, Elizabeth Graham designated the “green” city to refer to the Maya sites (and see Graham, 2006), a proposition further developed by (Isendahl and Smith, 2013).
It is therefore not tenable to argue that all settlements will have a simple spatial dichotomy between occupation area and crop resource hinterland, with the latter delimiting the extent of the former. When we add the information that the house gardens of Amazonian forest settlements provide a substantial portion of the domestic food supply (van der Waal, 2018) then the notion of what we mean by a hinterland for Angkor—and for other low-density, dispersed urban settlements needs to be robustly redefined, since houses in such settlements were presumably surrounded by economic trees and plants which would have made a substantial contribution to the domestic food supply.

A broader issue of taxonomy also needs to be addressed for low-density settlements. The terms we use for referring to settlements require rethinking. Formerly size, in terms of settlements area, could be a rough proxy for population and in stage theory settlement extent followed a linear trend from small camp to larger village to bigger town to huge city to colossal industrial megalopolis. No more. Greater Angkor covered the same area in the twelveth century as industrial Sydney did in Australia in 1945. And Angkor had bigger built reservoirs. But it was not an industrial city. Likewise, the low-density Giants may cover similar settlement areas as the conventionally defined, early urban settlements of China, Mesopotamia, Mesoamerica and the west coast or S America. And they can have large monuments. But that does not itself make them the same kind of places as the conventional defined compact, early cities, to all be designated under one label. We do, however, have to solve the conundrum of how to refer to the “Giants.” They cannot just be villages because that term is already used for numerous places in Europe and for the places like the Tallensi and Lobi Dagati settlements which consist of many domestic residential units—some larger than others—which do not have the equivalent of Cahokia’s Monks Mound or the Acropolis of Great Zimbabwe. This is a quandary, as the “Giants” are also not the same as Greater Angkor but are attractively labeled as dispersed, agrarian urban settlements. Low-density settlement patterns are a well-established, inherent feature of human behavior. What we need are suitable ways to refer to them which will facilitate our understanding of them and their long-term role in human community life.

CONCLUSIONS

Dispersed settlement forms and multiple trajectories of growth are a necessary part of models of urbanism. By including them we need to decide whether we will further extend and dilute the meaning of “urban” or will seek new, additional ways to specify what we are talking about. Terminologies are crucial to inquiry not only because they aid communication about a topic but also, conversely because they affect what we “see.” Labels are not neutral—they relate to theoretical premises and they in turn define how empirical data become meaningful. "Urbanism" as a label with diffuse meaning is liable to generate diffuse cross-comparisons which retain older theoretical premises. While urbanism and urban will necessarily continue to be used and will likely transmute to a new meaning, just as “gravity” did from Newtonian to Einsteinian physics, other foundations for cross-comparison can be usefully developed.

A agrarian low-density urbanism, the “Giants” and contemporary megalopi and desakota can be cross-compared in terms of form and trajectory. Because this frame of reference is spatial and quantified, it specifies no equivalence of sociality, though it does specify similarities in the demands and management of interaction and communication. In this comparison the consistencies are trajectories of similar “direction” but very different magnitudes. By contrast, if we use “urban” to subsume all these kinds of settlements and all the range of compact urban settlements as well, we do specify a resemblance of sociality because that is what “urban” has come to mean while also being very diffuse. In addition, that resemblance is a quality not a mundane quantity. Therefore, if we use the term “urban” we will need to rigorously specify qualifiers such as “industrial” and the already defined “agrarian, low-density urbanism” i.e., cases like Angkor, Tikal, Pagan and Anuradhapura and we will likewise have to specify an “urban qualifier” for the “Giants” and a substantial definition of the difference that it incorporates.

The materiality of settlement form in terms of spatial magnitude and internal density becomes recognizable as a factor in its own right; rather than as an epiphenomenon of sociality. Denser occupation areas, in any form of settlement, become cross-referable. Defining the internal density structure of settlements and the characteristics of the edges of settlements and designating the location and characteristics of hinterlands become critical issues for all settlements instead of mundane and self-evident features, as in the normalized perception of compact settlements.

If, as appears to be the case, the trajectories of dispersed, low-density settlements lead to very different outcomes then materiality has serious consequences and needs to be habitually incorporated into models of the formation, expansion and demise of human communities. The implications are potentially very serious, and the numerous cross-comparison offered by the past are therefore of great potential value as guides to what can happen and perhaps an indicator of the consequences of various social and material alternative for coping with the situation.

The three trajectories to low-density settlement form which can be referred to as the “Giants” trajectory (from 4000 BCE), the agrarian urban, low-density trajectory (from the forth century BCE) and the industrial-based conurbations, megalopi and desakota (from the late nineteenth century CE onwards), offer a new window into the history of settlement growth and the meanings of the term “urban.” They also offer a global basis for comparison and raise serious issues about the resilience of the low-density settlement form. Worryingly, the successive low-density settlements do not derive from the low-density cases of the preceding trajectory. Neither Angkor nor the Classic Maya cities had any connection to modern, industrial low-density cities. By contrast compact cities, the epitome of the conventional definition of cities, display continuity through successive urban settlement forms over several 1,000 years. The implications for modern, giant, low-density cities are ominous. They appear to face the risk of having no future.
AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This paper is funded via an OPEN-AIRE grant to Dr. Francesca Fulminante through the University of Rome (Marie Curie Sklodowska Fellowship 628818—Past-People-Nets 2014-2016).

ACKNOWLEDGMENTS

I am deeply grateful to Francesca Fulminante for inviting me to produce this paper. The task revealed far more than I had expected about the sources of my own awareness of low-density settlement patterns. To Meyer Fortes and especially to Jack Goody I will be forever grateful for the opportunities which took me to extraordinary places in Northern Ghana. To Christophe Pottier, the many members of the Greater Angkor Project and to the staff of APSARA and members of the government of Cambodia I am very grateful for the engagement, work and support which enabled the redefining of Angkor as a low-density city. To Luis Bettencourt, Scott Ortman, Mike Smith, Jose Lobos, and Sarah Klassen I am grateful for discussions about central areas of urban settlements, scaling, and the outer limits of settlement space. For illustrations in this paper my thanks to Jerry Sabloff for advice, to Bob Hudson, Scott Hawken, Todd Whitelaw and Andrew Wilson for images which they produced and to Kirrily White for preparing even more at very short notice. Much appreciated.

REFERENCES


**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Fletcher. *This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).* The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Modeling the Rise of the City: Early Urban Networks in Southern Italy

Lieve Donnellan*

School of Culture and Society, Aarhus University, Aarhus, Denmark

The rise of the state in Ancient Italy went hand in hand with an increase in infrastructural power, i.e., settlement centralization and urbanization. The paper discusses theoretical challenges and introduces a modeling approach to a case study, one of the earliest cities in Southern Italy, Pontecagnano, with the aim of understanding the community dynamics at the time of the earliest urbanization (ca. 900–600 BC). The model is a two-mode model that derives from social network analysis, an approach that has been fruitfully adapted to archaeological research. The model is applied to detect trends in burial contexts from the community involved. Burial was, at that time, in the region, a key instrument in the creation of memory and display of status and thus for building and consolidating state power. The analytical network model is able to detect the dynamics in the community over time very well: network Cohesion is expanding and contracting, and points to the existence of tension and a tight control of funerary behavior. The study of Centrality of selected nodes provides a good understanding of the strategies in terms of the circulation of key resources. The latter is particularly significant for studying urbanization because the appropriation of resources was not possible without centralization and the development of infrastructure, as well as an ideology. Based on the study of selected resources, it is suggested that an increase in crop storage has played a particular role in the development of state power and the urbanization process at Pontecagnano. In due course, the paper also addresses methodological challenges of working with fragmented datasets when applying models to study the past.

Keywords: modeling, ancient cities, urbanization, archaeology, Southern Italy, network theory

INTRODUCTION

The appearance of cities sparked a process of deep transformation in the Ancient World. The Ancient Cities of the first millennium BC testify to a dramatic change in economic integration, social interaction and political complexity—a leap forward on a global scale which was never fully undone. The philosopher Karl Jaspers coined the term “Axial Age” to indicate this evolution which appeared more or less simultaneously in China, India, Persia, and Europe (Jaspers, 1949).

The first scholar to write extensively about the Ancient City, Fustel de Coulangé (1980), is indicative for the general approach to urbanization as a historical process: “... just as several phratries were united in a tribe, several tribes might associate together, on the condition that the religion of each should be respected. The day on which this alliance took place the city existed. It is of little account to seek the cause which determined several neighboring tribes to unite” (Fustel de Coulangé, 1980, p. 119). Fustel de Coulangé considered religion to be the binding force of the Ancient City, independent of whether the tribes united voluntarily or were
coerced by an external force. Cities underwent a series of transformations, in his view, as the result of lower social classes (the plebeians) demanding the right of involvement in political government. The “revolutions” he identified as the motor behind social change, did occur in all ancient cities, but not at the same time.

Archaeologists and historians did not really pick up on the theoretical challenges of defining the Ancient City, but Fustel de Coulanges was very influential for later conceptualizations of cities and society among sociologists, through his most famous student Emile Durkheim. Other social thinkers, especially those interested in the city and urban phenomena, such as Weber (1921) in his posthumously published work discussed extensively the conditions of Ancient Cities such as Rome or Athens (among others). Likewise, the Ancient City occupies a central place among prominent early scholars of Urban Studies such as Simmel (1903) and Spengler (1922). Their studies, however, were based on the consensus of the day and did not contribute to a critical re-evaluation of Ancient Cities or urbanization processes among historians and archaeologists. Even today, despite the long history of interdisciplinary interaction and the comparative value that Ancient Cities bring to Modern Cities and vice versa, little is done to bridge the gap between sociology and the historical sciences.

Only in 1950 did an extensive comparative study by Childe result in a critical evaluation of archaeological and historical aspects of the urban character of settlements (Childe, 1950). Childe advocated the use of a checklist approach to distinguish cities from other settlements and he heavily favored monuments, law, writing, the market for exchange, and high culture as identifying elements. This checklist approach remained in favor for many decades among archaeologists and only more recently have more subtle ways of tackling the question of the nature and coming into being of Ancient Cities been proposed for the Mediterranean (Damgaard Andersen, 1997; Osborne and Cunliffe, 2005). Mediterranean urbanization, it was thus proposed, should be seen as a complex process of social, economic and political transformation in which two parallel dynamics are at work: state formation and settlement nucleation.

The adoption of new methods in archaeology in the late 1950s and 1960s, in which fields are walked systematically to record ceramic distributions, has revolutionized the study of Ancient Cities. Both in Greece and in Italy, settlement systems began to be studied in a regional perspective, which allowed to distinguish settlement hierarchies and transformations of settlement systems through time (Ward Perkins, 1961; Bintliff, 1999, 2000, 2002, 2006, 2016; Stoddart, 1999; Peroni, 2000; Pacciarelli, 2001; Bintliff et al., 2017). Urbanization came to be studied from a long-term perspective and within a broader framework of ancient landscapes.

An important consequence of this long-term perspective in Italy was, moreover, the realization that the earliest urbanization predated the appearance of Greek cities on the coasts of Southern Italy and Sicily. Up until recently, scholars had heavily debated the role that the urban culture of the Greeks and Phoenicians had on Italian societies, particularly on Rome and the Etruscans (Peroni, 2000; Riva and Vella, 2006; Riva, 2010; Fulminante, 2014; d’Agostino and Gastaldi, 2016). But the roots of Italian urbanization seem to be firmly placed in the Final Bronze Age transition to the Early Iron Age—the eleventh/tenth centuries BC (di Gennaro, 1986; Guidi, 1998; Peroni, 2000; Pacciarelli, 2001; Fulminante, 2014).

One of the major challenges in studying urbanization in Italy is the lack of physical remains of the earliest cities. They lie buried under modern cities or have suffered significant destruction throughout the millennia. Intensive archaeological analysis, including the collection of ceramics from the surface and excavations, have resulted in a general understanding of the spatial development of the earliest urbanization processes. In Central Italy, dispersed pottery scatters have been documented on large plateaus (from 20–30 ha up to 80 ha), where the later Etruscan and Latin cities were located (Pacciarelli, 2001). Initially, the pottery scatters were seen as belonging to pre-urban, small and distinct, settlement nuclei, that after ca. two centuries came together in a process of aggregation (synoecism), to form an urban settlement (Ward Perkins, 1961). This process of aggregation has often been attributed to historically known hero-founders such as Romulus in Rome or Theseus in Athens. Even though steeped in legend, many scholars today still believe that there is a ground of truth in these founder-figures (Carandini, 2018).

Around the plateaus on which the settlements were located, burials plots have also been found (Pacciarelli, 2001; Fulminante, 2014). Initially, these were seen as belonging to the pottery scatters on the plateaus, i.e., separate nuclei (Ward Perkins, 1961). Now, it is believed that the settlement nuclei, including burial plots, were more integrated than previously thought. Consequently, it was suggested that the different burial plots might rather belong to different social classes, political groups or other social divisions within society (Fulminante, 2014, p. 8–9).

The short physical distance between the settlement nuclei means that arrangements, e.g., about the use of space, field boundaries, etc., must have existed before the settlement was fully centralized. The development of the state and urbanization thus, must have been one of transformation of power and the use of space, i.e., integration and centralization over time, rather than a sudden event and radical transformation in terms of a foundation ex nihilo.

The last decades, archaeologists have focused on diverging trajectories of urbanization and the underlying social and economic transformations that may have caused settlement centralization (Damgaard Andersen, 1997; Nijboer, 1997; Pacciarelli, 2001; Osborne and Cunliffe, 2005; Motta and Terrenato, 2006; Guidi, 2008; Fulminante, 2014; Bintliff, 2016; Fernandez-Götz and Krausse, 2016). The establishment of cities is thought to have coincided with the consolidation of the state, although that centralized power can exist without a centralized settlement (Morgan, 2003; Osborne, 2005). Especially in Greece the phenomenon of non-urban centralization through federations or ethne has been well-studied (Morgan, 2003).

In Italy, the appearance of nucleated and centralized settlements, early “cities,” coincides with a marked rise of wealth deposited in tombs. It is thought that these tombs belong to a ruling “princely class” which consolidated its power and justified its position by adopting a new funerary ideology (d’Agostino,
How exactly these new elites exercised their power and came to be at the head of their communities is still unclear. The question as to why exactly the new state needed settlement centralization has also not yet been satisfactorily answered. Elite interaction with the Greeks and the accumulation of exotic objects could have been one strategy, but again, this explanation resorts to the “Greeks” as an explanatory factor in urbanization.

Thus, even though the general outlines of the urbanization process in Italy appear to be well-defined, several questions remain only partially answered. The mechanisms of the development of social and economic power underlying the urbanization process are still not well-understood. Yoffee sees state formation as a process of social differentiation and integration of the groups in a political framework (Yoffee, 2005). The way to understand early states, according to Yoffee, is by looking at interactions and tensions between different social groups and their leaders. Yoffee also points out that the interaction processes are complex and are not controlled by a single mechanism.

Yoffee himself rejects the possibility to model state formation processes mathematically, precisely for this reason of complexity (Yoffee, 2005, p. 169). However, as the present paper aims to demonstrate, a network-based exploratory approach is most appropriate to study early state formation and urbanization. Exploratory network analysis operates with a model to analyze complex datasets and, thus, provides a bottom-up approach to explore real-world data. The model is fairly simple and deeply embedded in social theory. Rather than taking an a priori emergent property at its core, the analysis tries to reveal markers of a process of diversification. The suspected complexity of the processes necessitate such a basic model.

Indeed, as Yoffee and others (Pacciarelli, 2001; Vanzetti, 2002; Fulminante, 2014) point out, the rise of the earliest cities was accompanied by social tensions and differentiation processes. The lack of written sources and the fragmentary archaeological data make that we have no information about the development of state power and elite agency in Ancient Italy. Inequality in this period seems to have been expressed in an archaeologically visible way in burial. Burial was in this time one of the main—if not the main—contexts for the creation of memory, the construction of social and economic differences and the negotiation of political power (Cuozzo, 2003; Fulminante, 2003; Laneri, 2007). We are unaware of the existence of other contexts in which inequality and power were expressed, e.g., sumptuous living, luxury dress or the consumption of exclusive food and drink, ritual activity. These contexts should not be excluded, but remain, at present, archaeologically understudied for the region and period in question.

Analyzing burials is, therefore, the key to studying the social tensions that scholars identify as underlying state formation processes, and ultimately, urbanization. Studying burials and social differentiation is, however, complex, because of the large quantities of data involved (Fulminante, 2003; Nizzo, 2015). From the well-studied early urbanizing communities throughout Italy, come hundreds, sometimes even thousands of tombs. Quantitative methods, combined with qualitative analysis, are therefore, of fundamental importance for the study of burials.

One of our best known sequences of burials of an early urbanizing community from Iron Age Italy comes from the South Italian city of Pontecagnano (Figure 1). Even though more fragmentary for some stages, Pontecagnano provides an exceptional source of information, not in the least for its extensive and detailed state of publication. With some notable exceptions for the later eighth and seventh centuries BC (Cuozzo, 2003), past research on the burials at Pontecagnano has been qualitative, rather than quantitative. Pontecagnano is thus particularly suited for the testing of an analytical model geared toward studying the social dynamics that underlie urbanization.

The analysis of burial data described in the next paragraphs demonstrates that, in contrast to what qualitative analysis often seems to suggest, the expression of social differentiation in burial in the urbanizing communities in Italy was not a linear process moving from simple to complex. Variations in Cohesion metrics over time point out that, at times when there was less quantity in objects deposited in tombs, special effort was placed on quality (diversity e.g., exclusive and exotic objects, or new spatial manipulations). With the use of a model it is possible to identify these phases of expansion and contraction, a process which is highly indicative of social tension between different interest groups in society. A regulatory body, reminiscent of an early state invested with politico-religious power must have been in place to oversee the burial process, i.e., the creation of social memory and the exercising of power through the expression of status and wealth.

The model also enables to focus on more detailed aspects. The study of the circulation of selected objects through Network Centrality values provides an insight in strategies of manipulation of resources by elites. The most important trend that the analysis picked up is the increasing role of storage vessels in the tombs at the time of supposed settlement centralization. Collection and storage necessitate a certain level of control over the population and an appropriate infrastructure. The collection and storage of agricultural surpluses are therefore well-known as one of the main drives behind centralization and urbanization, and indicative of what the sociologist Mann (1984) calls increased “infrastructural power” of the state. The increased emphasis on storage vessels in burial is highly suggestive for a rise in importance of the collection and storage of agricultural surpluses in real life, as burial appears to have been the arena where status and display were increasingly played out (Cuozzo, 2003; Fulminante, 2003). The study thus proposes that, at Pontecagnano, collection and storage of agricultural surpluses was an important factor in urbanization.

**MODELING ANCIENT CITIES**

In contrast to modern cities, Ancient Cities have not received an overwhelming attention from a modeling perspective. Archaeological approaches to the earliest cities are very often merely descriptive, aiming at the classification of objects and structures in terms of type and chronology, or, at locating the
remains of buildings on a map. Whereas, typo-chronologies can be very useful as a collecting strategy, they do not provide solid explanatory frameworks. Often, the use of typo-chronologies results in outdated culture-historical narratives and a reliance on much later written sources. Yet, there is a huge potential for studying ancient cities with computed models. Modeling allows to propose hypothetical reconstructions of fragmentary data, it contributes toward testing hypotheses and enables the discovery of patterns in large and complex datasets. Surprisingly, archaeologists have not adopted modeling approaches on a wide scale, despite the many advantages it may offer in complementing qualitative archaeological analysis. Models, such as the one used for the analysis in the present paper, need not be overly complex and can be heavily theoretically informed, yet very powerful as a heuristic tool.

Since decades, archaeologists have used modeling for all kinds of different questions, usually to study space, e.g., models for predicting settlement location (Bevan and Lake, 2013; Verhagen, 2018). Other models have focused on estimating settlement boundaries, e.g., with the XTENT model (Ducke and Kroefges, 2008), or explored the exploitation of ancient territories (Farinetti, 2009). Fruitful modeling has been applied to inter-visibility (Brughmans and Brandes, 2017) and the reconstruction of ancient transportation networks (Groenhuijzen and Verhagen, 2017).

Recently, modeling approaches have yielded positive results in the study of urbanization in Central Italy between the Bronze and Iron Age. In a series of studies Francesca Fulminante (2012a,b, 2014) (Fulminante et al., 2017) applied a network model to test various hypotheses regarding the formation of urban centers. By looking at settlements, locations and hierarchies as a network system, she formulated and tested a number of hypotheses regarding growth. One of the most striking conclusions that came out of the studies was the existence of regional differentiation in urban growth. In the region of Latium Vetus (later: Rome) the rich-get-richer effect seems to have subtracted increasingly people and resources from the surrounding area (Fulminante and Stoddart, 2010, 2012; Fulminante, 2012a,b).

The latter, as well as some other studies (di Gennaro, 1982; Guidi, 1985; Redhouse and Stoddart, 2011), rely on spatially determined models of urban settlement or growth (Gottdiener et al., 2005; p. 83–86; Li and Gong, 2016). Spatially-oriented views
allow to formulate hypotheses regarding the underlying social, political and economic dynamics at a macro level. These are extremely useful to develop broad historical reconstructions in the absence of written sources. More challenging with a modeling approach, however, is the study of dynamics at a micro or meso level. Usually, archaeologists take settlements and regions as unit of analysis, and operate with datasets that, in fact, span centuries. With such an approach, it is difficult to obtain a finer chronological resolution or address agency. It is precisely here that an exploratory network approach, such as the one introduced in the present study, proves its utility: it uses elaborate datasets and a model developed to study human interaction at a meso level.

Recent advances made in adopting and adapting network-theoretic approaches in archaeology (Brughmans, 2010, 2012a,b; Knappett, 2011, 2013; Leidwanger et al., 2014; Collar et al., 2015), have enabled a whole new perspective on past human interaction. Following general trends in modeling in archaeology outlined earlier, the first examples of network analysis by archaeologists can be characterized first and foremost as spatial in nature. The previously cited studies of transportation networks (Groenhuijzen and Verhagen, 2017), visibility networks (Brughmans and Brandes, 2017), fall into this group, as well as studies of spatial organization of territories (Rivers et al., 2013a,b), or regional exchange systems (Mills et al., 2013, 2015).

Recent studies have applied network analysis to look at processes of social and cultural interaction and transformation such as the rise of ethnic identities (Collar, 2013; Blake, 2014; Peeples, 2018), to trace the diffusion of knowledge among artisans (Østborn and Gerding, 2015), or to study pottery production processes (Van Oyen, 2016). Most studies, however, depart from a site or region as unit of analysis. This scale of analysis is sometimes considered problematic for the perceived lack of agency (Knappett, 2011; Leidwanger et al., 2014; Collar et al., 2015; Van Oyen, 2016, 2017). Archaeological applications of network methods and theories yield, however, the potential to develop into a theoretically-enhanced approach to past materialized interaction (Donnellan, 2016a,b).

Focusing on communities of interaction may provide a fruitful path to explore the intersection between the social and spatial aspects of interaction at different scales (Peeples, 2018). Recent archaeological studies have devoted extensive attention to the study of communities and agency (Canuto and Yaeger, 2000; Mac Sweeney, 2011; Varien and Potter, 2011). The concept of “communities” also offers a firm theoretical foundation for the present analysis, as “communities” are a key concept in contemporary urban studies. The study of community formation and transformation within urbanizing processes are at the heart of urban sociological and geographical analysis (Gottdiener et al., 2005). The notion of “community” relates to Yoffee’s idea of archaeologically detectable social groups and their differentiation as basis for state formation (Yoffee, 2005). Moreover, community detection constitutes one of the corner stones of formal network analysis (Boishevan and Mitchell, 1973; Boishevan, 1979; Wellman and Berkowitz, 1988; Wasserman and Faust, 1994; Borgatti et al., 2013). The concept of communities thus allows to connect to a broad range of theoretical and methodological approaches.

Community detection was also at the heart of the study made by the sociologist Davis and his colleagues about race relations in the Deep South of the US (Davis et al., 1941). They studied the norms and behaviors that sustained the construction of racial and social identities among different groups of people. One of the groups of people they studied was a small group of upper class women that attended high society events. Based on the frequency of co-attendance of these events, Davis et al. concluded that these women formed subgroups or cliques. By interacting frequently on the occasion of the events, these women developed similar patterns of behavior and expectations. One of the central notions of network analysis is exactly this similarity in behavior of agents, “homophily” (Carrington and Scott, 2011; Borgatti et al., 2013). The tighter the network, the greater the similarity between actors. Innovations within the network, according to network theory, come from interactions with actors outside the closely-knit group, as the concepts of weak ties (Granovetter, 1973) and structural holes (Burt, 1995) indicate.

The Davis study is considered a “classic” example in formal network theory for the two mode model (Borgatti et al., 2013), despite the fact that Davis and his colleagues did not use formal network analysis, nor developed the well-known graphical representation with nodes and ties. The visual representation of social networks, the sociogram, had been developed a decade earlier by Moreno (1934), but was not yet applied widely at the time of the Davis study. Davis and his colleagues used a matrix to represent the frequency of attendance of the social events by the group of women (Figure 2).

As Davis demonstrated, focusing on similarities such as co-attendance of events is the key in detecting communities of interaction. Similar principles for community detection have been applied to study the network of intermarriage and business among the Florentine elite in the Renaissance period (Padgett and Ansell, 1993) and the notion of interlocking boards of directors (Mizruchi and Schwartz, 1992). In reality, all two mode network analysis takes this idea of interaction and community formation as starting point. The model and its theoretical foundation were therefore considered to provide a solid theoretical and methodological base and thus adapted to fit the question of community interaction and diversification at Pontecagnano.

In contrast to sociologists and historians, archaeologists do not have access to information regarding the attendance of events, intermarriage or business ties. However, archaeologists can focus on the material remains of interaction. Building styles, pottery production styles, burial rites, etc., present similarities and differences in production, appearance, and consumption and can thus be indicative of close interaction. Similarities in behavior, called “homophily” in sociology (McPherson et al., 2001), can be voluntarily or coerced, but is essentially present in one way or another in cliques or network clusters. Even though the original process of tie formation, e.g., co-attendance, marriage, or business cannot be detected by archaeologists, similarities in the material world can act as a proxy for social interaction and can be studied in terms of markers of community formation. Even though fragmentary in nature—called the black
box problem—general trends of community formation can indeed be detected in archaeological datasets (Sindbæk, 2013).

Obviously, not all close interaction results in material similarities, nor does material similarity necessarily indicate close interaction, as it can be simply a coincidence. Key is to relate the material record to identifiable social behavior and contexts of interaction. The way an archaeological interaction model thus operates is by documenting all material features in the archaeological record that relate to identifiable depositional practices, e.g., burial. This materialized dataset of coherent behaviors, in its entirety, constitutes a network. Ties between features are next created when the features share similarities, following the principle of the two mode model. The more ties, the closer the similarities and the more important the homophily between the agents.

In reality, there are several ways of constructing a model. For the model in the present paper, the two mode model of Davis was followed closely and adapted to the specific research question: Davis’ women are “translated” into tombs and the events into the material features of the burial (Figure 3). The material features are considered a proxy for the event: the whole series of gestures, vocal expressions, movements, and perishable objects, etc., that were used in burial (Nizzo, 2016). The key assumption is that similarities in material expression are indicative of close community interaction (be it voluntarily or coerced), and dissimilarities are evidence of diversification. The diversification processes, thought to underlie state formation as outlined supra, can thus be studied with what is, essentially, a similarity matrix.

The model—a standard two mode network model—can be manipulated for analysis using a number of algorithms (Borgatti and Everett, 1997; Opsahl, 2013). These analytical tools are included in several standard consumer software programs for network analysis, such as UCINet (Borgatti et al., 2002), which was also used for the analysis reported in this paper. There are many analytical procedures available. Relevant for the case study are: Cohesion and Centrality, whose definition and analysis are reported below.

THE CASE STUDY

Background

Pontecagnano is a well-known archaeological site, located in the region of Campania, province of Salerno, in Southern Italy and was one of the most important Southern Etruscan cities, whose name, however, is not known with any certainty. In 268 BC, the Romans re-founded the city as Picentia.

Emergency excavations in the context of large-scale infrastructural works were conducted by the local archaeological services from the 1960s onwards, with important contributions from the University of Salerno for the study and publication of the results. These excavations resulted in the discovery of various nuclei of tombs, located east and west of the city, as well as traces of urban architecture, mostly of later date (d’Agostino, 1968; Cuozzo, 2003; Cuozzo et al., 2004; Bonaudo et al., 2009; d’Agostino and Gastaldi, 2012, 2016; Pellegrino, 2015). The excavations allowed to propose a hypothetical reconstruction of the regular lay-out of the Roman city, which apparently overlays the earlier Etruscan city. The topography of the earlier phases is only known in part, however, the general spatial development of the settlement through time has been reconstructed (Figure 4).

The earliest remains that testify to human occupation are burials in the areas of Pagliarone (Gastaldi, 1998) and Picentino (d’Agostino and Gastaldi, 1988; Cinquantaquattro, 2001; De Natale, 2016), dated to the early ninth century BC. Through time, new plots were taken into use at San Antonio, in the later ninth century BC (De Natale, 1992). The Pagliarone plot does not offer extensive evidence beyond the earlier phases, whereas the other plots continued to be used for many generations. At the turn of the eighth to seventh centuries BC, there is evidence for an extensive reorganization of the funerary landscape, with new plots taken into use, e.g., at the Piazza Risorgimento (d’Agostino,
1968), whereas other plots such as on the contemporary Corso Italia (Cerchiai, 1987) or the plot of Casella (Cinquantaquattro, 2001) testify to new extensions to existing burial grounds. These reorganizations are considered to be the hallmark of the early urban community (d’Agostino, 1968; Cuozzo, 2003; Cuozzo et al., 2004; Bonaudo et al., 2009; d’Agostino and Gastaldi, 2012, 2016; Pellegrino, 2015).

Between the various plots and through time, there is a marked variation in burial practices that testify to the existence of dynamic communities and interaction, which will be the object of the analysis. The burials consist of simple pit burials for inhumation (*a fossa*) or for cremation (*a pozzo*). Slightly more elaborate architectural shapes exist already for inhumation, with the bottom, walls and cover dressed with stones (Figure 5). Cremated remains could be deposited in a pit, covered with a large stone (*a ricettacolo*), sometimes elaborated with a sort of platform on which objects could be displayed (*a vestibolo*). The typical way of depositing cremated remains, in a biconical urn covered with a one-handled bowl, or sometimes a ceramic helmet, echoes the practices known in Central Etruria, in the area of the so-called Villanova groups, considered the predecessors of the Etruscans (Bartoloni, 1989).

On the occasion of the burial, various objects were deposited simultaneously with the remains of the deceased. These objects...
could range from a variety of ceramic containers for the consumption of food or drinks, for pouring or storage (anfora, anforiscos, askos, piatto, scodella/-one, tazza, and other shapes), items for dress and bodily care and adornment (fibulae, rings, clasps, razors, etc.), equipment for weaving and a variety of weapons.

The material culture at Pontecagnano has been the object of sustained study by archaeologists, and as a result, our knowledge of the typological variety of the local material culture is decidedly among the best in the whole of Italy for this period (Gastaldi, 2016). Not all contexts have been published yet and several plots are still in course of analysis. Most importantly, the vast majority of the tombs of the so-called Orientalizing period (late eighth to seventh centuries BC) are currently still in course of study. The future publication of new information will provide an opportunity to test the hypotheses formulated in this paper. Some plots have been published, such as at the Piazza Risorgimento (d’Agostino, 1968), Corso Italia (Cerchiai, 1987), as well as in the above-mentioned burial sites of Picentino at Località Casella (Cinquantaquattro, 2001) and San Antonio (De Natale, 1992), but some plots during certain time intervals have given evidence of only one or two burials (S. Antonio I-IB; Piazza Risorgimento; Corso Italia). These plots have not been included in the quantification as network analysis does not function well with very small amounts of data. Future analysis could focus on the integration of the different datasets into one larger dataset, including the material that is still being studied.

Importantly, a study by Cuozzo (2003) has dedicated ample attention to the unpublished Orientalizing tombs and reconstructed the social and political transformations of the period. Cuozzo applied a cluster analysis to trace the dramatic increase in social differentiation in the late eighth and the seventh centuries BC. Social differentiation was much less pronounced in the earlier Early Iron Age phases, and Cuozzo explained the phenomenon as a conscious strategy to support an ideology of power, carved out by a new, urban, political elite. Part of the elite strategy was a transformation in the use of space through the relocation of burial and settlement centralization. Cuozzo’s study has been particularly important for providing tangible evidence for socio-political dynamics underlying the general patterns of urbanization identified in other Central-Italian centers.

However, as was explained previously, despite the major achievements of existing studies, several questions remain unanswered. The earlier phases at Pontecagnano remain understudied in quantitative terms and a quantitative comparative analysis of the different plots, all published in different volumes, has not been undertaken. It was outlined supra that the exact mechanisms of integration of the various nuclei into a single urban community are not described in detail and the diversification of the community within the process of state formation has also not been defined in detail. By using a two mode model, or similarity matrix, the analysis discussed below aims at describing the diversification process in more detail.

**Analysis of the Case Study Datasets**

As a first step, the analysis started with manually digitizing the data of published archaeological contexts (tombs). The dataset was collected and stored in Access® to allow an easy extraction to compile datasets in the UCINet format. The information collected in the database concerned material, shape, type, date and context of objects and rite, tomb shape, date, and location for context. Additional information such as orientation, gender, age was collected for tombs as well. Detailed typological
classifications of the objects as defined by the excavators (Gastaldi, 2016) were also inserted in the database although the present study will not take this level of detail into consideration. The focus lies on object shapes, e.g., cup, bowl, or fibula, etc. A study of the circulation of the detailed object typo-chronology defined by the excavators would, no doubt, reveal new dynamics, but this would ideally be undertaken together with an archaeometric analysis, as typological classification does not necessarily coincide with production units. An analysis purely based on typological units would thus, not necessarily provide all the details one would, ideally, wish for.

An issue in the analysis is certainly the integrity of the data, as not all tombs are well-preserved. Some tombs were destroyed already in Antiquity, others more recently. Tombs or objects that were only partially preserved have been omitted from the analysis if they were beyond basic recognition, i.e., shape. Another issue is that not all tombs are well-dated. The analysis relies on the traditional chronological classification in phases. The tombs for which only a very broad chronological estimate could be given, e.g., a date within the whole “Early Iron Age” (which spans almost two centuries!), have also been omitted from the analysis. Tombs that have been dated tentatively to a chronological range, e.g., Early Iron Age I rather than phase IA or IB, have been included twice in the analysis, i.e., both in phase IA and IB.

There are obviously several ways to deal with impartial and fragmentary data in archaeological network analysis. Ideal would be to develop different approaches, varying from more detailed, omitting uncertainties, to broader, including the more fragmentary evidence. The various datasets could then be compared to assess which one provides the better results. However, limits posed to space in this paper do not allow such an elaboration, although future analysis should certainly aim to address the factor of uncertainty in a more elaborate way.

A series of datasets was compiled for every published burial nucleus, per time slice (chronological phase). Thus, datasets were developed for Picentino Phase I-IA (c. 900-850 BC), I-IB (c. 850-780/70 BC), II-IIA (c. 780/70-750 BC), II-IIB (c. 750-730/20 BC), Pagliarone I-IA (c. 900-850 BC), I-IB (c. 850-780/70 BC), San Antonio II-IIA (c. 780/70-750 BC), II-IIB (c. 750-730/20 BC), c. 730/20-675 BC, c. 675-600 BC, and Loc. Casella c. 725-700 BC. The other plots with only one or two tombs were excluded as small numbers do not allow for network modeling, as was outlined supra. The results reported below were obtained via standard exploratory analysis of Cohesion and Centrality using UCINet (Borgatti et al., 2002).

Cohesion

Network cohesion can be understood as measure of how closely connected or tightly knit a network is, i.e., in this specific case study: how similar or dissimilar the burials are. Network cohesion can be calculated in virtually any network analysis software package, using standard algorithms. The calculation gives a relative value for the network as a whole (Borgatti et al., 2013): if everyone knows everyone, cohesion is 1.00.

The cohesion measure includes the calculation of a number of elements (Borgatti et al., 2013), such as density or degree (the number of ties), and distance between the nodes: the minimum distance (radius), maximum distance (diameter) and average distance. Average distance indicates how many steps must be taken on average to arrive from a node to a randomly chosen other node, traveling via shortest paths (Borgatti et al., 2013). Diameter gives the longest of all shortest paths in the network, whereas radius is exactly the opposite, and gives the shortest of all paths (Borgatti et al., 2013). There are other measures to calculate cohesion, but these will not be discussed further within the context of this paper, as they do not provide any added value to the discussion.

The Cohesion metrics for the datasets have been calculated in UCINet and the results have been plotted in a graph, to enhance readability (Figure 6). The plot showing density (a) clearly shows variation in the density, i.e., the number of ties, or objects associated with tombs. It is often easily assumed that the diversification process that accompanied state formation and urbanization in Italy was a linear one, in which a group of people gradually deposited more and more objects in the tombs. The graph clearly shows that this was not the case and that the average quantity actually drops (phases I-IB and II-IIA in Picentino) before it rises again, in the S. Antonio and Casella plots after phase II-IIA.

The other Cohesion measures focus more on similarity and dissimilarity. The lower the distances between the nodes, the more similar the tombs were in terms of objects deposited. Radius provides a value for the minimum number of steps that have to be taken from one node to reach another (this would be calculated in terms of co-association of objects in the tombs: through co-association, one could virtually travel from one node to the other). Diameter gives a value for the maximum number of steps that have to be taken (again through the virtual steps of co-association of objects in the tombs) and average distance provides the average value.

The patterns revealed by the graph are highly interesting as (1) they also do not show a linear development from more similarity to less similarity through time and (2) in conjunction with the values for degree, they clearly show the existence of tensions in the system.

A first rise in diameter can be observed in Picentino in phase II-IIA. This indicates that there is an increase in dissimilarity in terms of objects deposited in tombs (a larger variety of types). This trend is at odds with the low value for degree: at exactly the same time, we can observe the lowest degree value for the whole network. This suggests that, while people deposited fewer objects in the tombs, they sought out larger variation. This trend is highly suggestive for the existence of social tensions, and possibly, limits that had been enforced on the spending in burial rites: one could not deposit too many objects at the time of burial.

The other striking trend at S. Antonio for the same period reveals the opposite: people deposited more objects at the time of burial, but the objects were less diverse: the focus was on quantity rather than quality. Through time, this pattern continues to exist at S. Antonio: the number of objects deposited in the tombs increases, but there is not so much diversity. Again, this can be understood as a measure of reinforcing certain behaviors in burial, be it self-imposed or coerced, in which ostentatious depositing of wealth was not permitted.
A next series of measures that was performed on the datasets concern Centrality. Whereas, Cohesion looks at the network as a whole, Centrality operates at node level and assesses the degree of importance of a node within the network (Borgatti et al., 2013).

There are different ways to calculate Centrality. Degree Centrality is calculated based on the number of vertices incident to the node, i.e., the number of ties a node has. A variation of Degree Centrality is Eigenvector Centrality. This measure counts the number of nodes adjacent to a given node (just like Degree Centrality), but weights each adjacent node by its Centrality (Borgatti et al., 2013). Other measures of Centrality that can be calculated are Closeness, which is based on the sum of the geodesic distance (i.e., the length of the shortest path) from a node to all other nodes. Betweenness centrality indicates how often a given node falls along the shortest path between two other nodes (Borgatti et al., 2013). In terms of our model, a high Degree Centrality means that an object occurs frequently in the tombs, whereas the measures of Closeness Centrality and Betweenness Centrality indicate how often it is associated with other “popular” objects. This measure, in fact, looks to what extent a node can be considered part of the “norm” in funerary behavior at a specific time.

Because of limited space, the present paper focuses on describing the circulation of a selected number of objects via Centrality measures only. Future studies can focus on...
other material categories, as well as on individual tombs. The object groups whose circulation is discussed are: vessels for the consumption of food and drink (scodella and tazza), vessels for the pouring, consumption and possibly libation of liquids (askos, brocca, attingitoio), status objects (the fusaiola or spindle whorl for weaving equipment and weaponry such as the ceramic helmet or elmo fittile, lancia, spade, and giavelotto), objects for bodily adornment (fibula and bracciale), and storage vessels (olla and amphora). Obviously, the archaeological record has revealed many more objects, but the selected objects are some of the most frequently encountered and can be seen as representative for general trends in the community.

The graph (Figure 7) shows that both cups and bowls (tazza and scodella) enjoyed popularity during the earlier phases, but declined after ca. the mid-eighth century BC, both in quantity (Degree) and relative importance (Eigenvector, Closeness and Betweenness). The increasing popularity of Greek-style painted fine wares at Pontecagnano (not included in the graphs) can be seen as an explanation: the new style of vessels were probably considered to be more attractive and they could easily serve as substitutes for the plumper handmade vessels of the earlier local tradition.

Among other vessels for transferring liquids (from one container to another, or to the ground in an act of libation or for consumption), two types see a sharp decline through time: the askos and brocca see hardly or no use in the later phases, as the graph (Figure 8) shows. However, despite the steep decline in numbers of the askos (lower degree value), the Closeness Centrality in Picentino II-HIB is still high. This indicates that what we have was very much part of the core of the network, suggestive for high similarity with the other tombs and thus, normative behavior.

Interesting is the increasing popularity of the attingitoio, a larder-shaped vessel. Whereas, the other vessel types might have been abandoned for more appealing painted fine wares, the attingitoio appears to have continued in use, despite being plump, unpainted and handmade. A similar development was observed in another (unrelated to this) study in the North Aegean in the Iron Age, where pouring vessels deposited in burials continued to be handmade, despite the availability of wheel made fine ware alternatives (Donnellan, 2017). A possible

---

**FIGURE 7** | Graph displaying Centrality values for selected objects related to the consumption of food and drinks (tazze, scodelle) in the analyzed plots per period: (A) bowls (scodelle) Degree, Eigenvector, and Closeness Centrality (scale %), (B) Bowls (scodelle) Betweenness Centrality (scale %), (C) cups (tazze) Degree, Eigenvector and Closeness Centrality (scale %), (D) cups (tazze) Betweenness Centrality (scale %).
Donnellan Modeling the Rise of the City

FIGURE 8 | Graph displaying Centrality values for selected objects related to transferring liquids (pouring/consumption): (A) askos Degree, Eigenvector, and Closeness Centrality (scale %), (B) askos Betweenness Centrality (scale %), (C) jug (brocca) Degree, Eigenvector, and Closeness Centrality (scale %), (D) jug (brocca) Betweenness Centrality (scale %), (E) larder (attingitoio) Degree, Eigenvector, and Closeness Centrality (scale %), (F) larder (attingitoio) Betweenness Centrality (scale %).

The graph looking at status objects (Figure 9) shows a decline through time for the spindle whorls (fusaiola). This is curious, as the role of textiles in the urbanization process has been underlined recently (Gleba, 2015). Despite the supposed importance of textile production and its aristocratic connotations, this was not expressed in the later tombs. Weaving was essentially a gender-related activity and it is possible that the expression of gender-related status was deemed less central in

explanation in the Aegean case was that the production of the vessels was invested within special, possible ritual, significance in which the vessels were attributed agency, which made that the production and consumption prohibited the use of mass-produced wheel made vessels. It can be hypothesized that the use of the attingitoio at Pontecagnano was attributed similar ritual agency, and therefore, the shape continued to be produced in a traditional way.
FIGURE 9 | Graph displaying Centrality values for selected objects related to status display: (A) spindle whorl (fusaiola) Degree, Eigenvector, and Closeness Centrality (scale %00), (B) spindle whorl (fusaiola) Betweenness Centrality (scale %100), (C) ceramic helmet (elmo fittile) Degree, Eigenvector, and Closeness Centrality (scale %00), (D) ceramic helmet (elmo fittile) Betweenness Centrality (scale %100), (E) lance (lancia) Degree, Eigenvector, and Closeness Centrality (scale %00), (F) lance (lancia) Betweenness Centrality (scale %100), (G) sword (spada) Degree, Eigenvector, and Closeness Centrality (scale %00), (H) sword (spada) Betweenness Centrality (scale %100), (I) spear (giavelotto) Degree, Eigenvector, and Closeness Centrality (scale %00), (J) spear (giavelotto) Betweenness Centrality (scale %100).
burial at the time of intensified urbanization in the late eighth to seventh centuries BC.

Among the weapons, a decline can be observed in the graph (Figure 9) for the ceramic helmets and swords (elmo fittile and spada). The ceramic helmets served no purpose in daily life, but are considered imitations of real bronze helmets (Egg, 2017, p. 167). Possibly, these were part of traditional forms of expression of role and status, together with the sword, whereas the lances and spears (lancia and giavelotto) (Figure 9) remain fairly constant in use through time. Moreover, the high Closeness values of the latter seem to indicate that they continued to occupy a central position in the network. This can be understood as the integration of these items into new urban elite behaviors. Spears are associated with hoplite armor, which is considered an urban form of organization of community defense, attested also later in Etruria (Egg, 2017). Such an organization of military forces can thus be seen as a clear expression of organized state power and urbanization and also of effective state formation.

Among the metal objects for dress and bodily adornment, the analysis looked at fibulae and bracelets. A large variety of metal objects have been found in the tombs at Pontecagnano, but both selected objects types allow to document the general trends in the network. Given the advance of elite power through time, the analysis was particularly interested to see whether the general access to metal objects would be restricted to elites as they grew more powerful with the advent of state formation and urbanization. As the graph displaying the metrics shows (Figure 10), the circulation of fibulae remains fairly constant through time, in terms of Degree, Eigenvector, and Closeness Centrality. The Betweenness Centrality values (Figure 10) display more of a decline, however. Betweenness indicates the role a node plays as a bridge in the connection of other nodes. The decline in Betweenness Centrality in our model could indicate that the fibulae did indeed become a little more confined to a group of tombs that were more part of the core, i.e., similar in content. The patterns in the graph are, however, suggestive for the traditional patterns of redistribution of metals remaining intact through time, at least until the late seventh century BC, despite the major transformations at a political level and the advance of state formation and urbanization. The deposition of bracelets even displays an increase, as the graph indicates (Figure 10). This can be understood as part of the general trend in the seventh century BC of depositing more wealth in the tombs.

The last group of objects whose circulation this paper will discuss are storage vessels. Ollae and amphorae are common household vessels that were repurposed for burial. The graphs (Figure 11) indicate that, initially, they were not very popular in the tombs. However, the olla sees a steep rise especially from phase II-IIA onwards, to decline again toward the late eighth century BC. The shape is not abandoned fully and even becomes slightly more popular again in the seventh century BC. The decline of the ollae seems to coincide with the increase in popularity of the amphorae. These see a gradual rise initially, but the development is much more outspoken after ca. 750 BC. This is an important development both in terms of state formation and urbanization.

The increased emphasis on the depositing of amphorae in the tombs comes at a time that a local production of amphorae was established on the nearby island of Pithekoussai, with which Pontecagnano interacted intensively. The Pithekoussan amphorae are of an entirely different type, but were essentially storage vessels presumably for the local wine production (Donnellan, 2019). The Pithekoussan amphorae were often repurposed in the local necropolis, almost exclusively for infant burials. The intensification of the wine production at Pithekoussai and the facilities needed for production, collection and circulation are thought to have been an important factor in the local urbanization process (Donnellan, 2019). It is very tempting to see an analogous development at Pontecagnano, even though it is unclear which production the amphorae might have served. In contrast to Pithekoussai, the amphorae in the necropolis at Pontecagnano were not used for the burial of infants. They were much smaller and would have been more difficult to reuse in a similar way as at Pithekoussai. The ideology behind depositing storage vessels, however, can be considered the same: signaling agricultural wealth and abundance. In Athens, most famously, one had to produce at least 500 medimnoi (an ancient unit of volume) of grain to belong to the highest classes and obtain political rights (Arist. Ath. Pol. 7.3-4). The identification of this trend at Pontecagnano is, therefore, highly significant for the social and political changes in the community.

**CONCLUSION**

Through the systematic analysis of burial contexts of the ancient city of Pontecagnano, it was attempted to gain a better understanding of the urbanization processes in Ancient Italy. Urbanization in Italy went hand in hand with state formation: settlement centralization occurred at the time of the establishment of a political class of rulers who signaled wealth and sustained power via ostentatious burial rites. This process has been observed in several Italian cities, all dated around the late eighth to seventh centuries BC. The cities started as a collection of settlement and burial nuclei, whose integration has been disputed. It is now, however, thought that a minimum of integration must have existed before settlement centralization (Fulminante, 2014).

Even though the broad outline of the urbanization phenomenon had been defined (a process of settlement centralization), many questions remained still unanswered. In particular, the dynamics at a micro and meso level were not well-understood: how state power developed and how state power could have influenced settlement centralization and urbanization. It was outlined how archaeologists consider state power as the result of a process of social differentiation and power struggles between interest groups in communities. It was explained how, in absence of other archaeological traces, burials provide the key to understanding processes
FIGURE 10 | Graph displaying Centrality values for selected metal objects related to bodily adornment: (A) fibula Degree, Eigenvector, and Closeness Centrality (scale %occ), (B) fibula Betweenness Centrality (scale %occ), (C) bracelet (bracciale) Degree, Eigenvector, and Closeness Centrality (scale %occ), (D) bracelet (bracciale) Betweenness Centrality (scale %occ).

FIGURE 11 | Graph displaying Centrality values for selected vessels related to food storage: (A) olla Degree, Eigenvector, and Closeness Centrality (scale %occ), (B) olla Betweenness Centrality (scale %occ), (C) amphora (anfora) Degree, Eigenvector, and Closeness Centrality (scale %occ), (D) amphora (anfora) Betweenness Centrality (scale %occ).
of social differentiation. As burial might have been one of the most, if not the most, important context for the creation of collective memory, expression of power and political negotiation in the region at that time, they are a particularly rich source to study state formation, and ultimately, urbanization.

The analysis departed from a model adapted from social network analysis: the two mode model or affiliation model (Borgatti et al., 2013). The model was initially developed to study social interaction and the formation of communities through shared interaction in the Deep South of the US (Davis et al., 1941). The general assumption of the study and, more generally, in network theory, is that (1) people who interact closely together form communities and share various behaviors and norms, (2) behaviors can be plotted in similarity matrices and graphs to provide a “tool to think with.”

The notion of “community” was considered to provide a solid theoretical base both in terms of urban studies and network theory. The Davis model for community detection and interaction was adapted to fit an archaeological reality in which material patterns are seen as a proxy for human interaction: material similarities or “homophily” are considered a sign of cohesion in the community, dissimilar patterns as a sign of diversification.

With the two mode model, networks for different datasets were constructed. The datasets stem from excavation data of different burial plots over time. The networks, essentially similarity matrices, were further analyzed using a standard network analysis software, UCINet.

The analysis revealed firstly that the development of diversification in burial was not a linear process. In fact, in the later ninth and earlier eighth century BC, there was either a decline in the minimum and maximum number of objects that were deposited in the tombs, and if there was not so much a decline in differentiation through numbers, then there was a lack of diversity in the types of objects placed in the tombs. These patterns can be considered highly significant for the existence of an authority who controlled burial practices and who imposed rules that agents tried to circumvent through innovation. For example, T 2145 of the Picentino necropolis of phase I-IB was a tomb with one of the highest degree values for that plot at that time (it contained 17 objects, including weapons and a large storage vessel). In terms of content and diversity, the tomb was not radically different from others. But, space had been manipulated: the tomb was located at the center of a group of monumentalized tombs, while it was itself monumentalized with a platform that was shaped like a horse shoe (d’Agostino and Gastaldi, 1988). This innovation in the use of space had not existed before, and it can be suggested that new solutions were sought for status display, in other ways than in depositing objects. These sorts of tensions are highly significant for the study of diversification within the community and the existence of different strategies for social competition. Patterns like these can only be revealed by looking at global network metrics of Cohesion, in addition to local measures of Centrality—combined with qualitative analysis of the use of space.

The study of the circulation of selected object categories revealed several interesting and highly significant trends for state formation and urbanization. The analysis registered the decline of several traditional pottery shapes (tazza, scodella, brocca, askos) that found easy substitution with new painted fine wares that were imported from elsewhere after ca. the mid-eighth century BC. The traditional handmade attingitoio continued to be used and even increased in popularity, suggestive for a continuity in certain traditional practices that might have included ritual agency attributed to this particular vessel shape.

Traditional forms of gender-related status such as weaving equipment, swords and ceramic helmets also declined over time. Dramatic social changes and changing social values accompanied the rise of the state and advance of urbanization and led to the abandonment of certain traditional practices. The analysis, in contrast, registered the continuity in the use of lances and spears. This is a significant development that hints at the adoption of hoplite warfare, a phenomenon that scholars have (controversially) sought to relate to the establishment of the polis in the Greek world (Viggiano, 2013).

The last significant trend that the analysis revealed was the increased emphasis on storage vessels in the tombs after c. the mid-eighth century BC. Storage vessels signaled agricultural wealth, the latter being a condition for citizenship and political rights in most Ancient agrarian states. Moreover, the collection, storage and redistribution (or repurposing) of agricultural surpluses required an increased control over populations and appropriate facilities. Thus, settlement centralization and increased state power were often fueled by this process. Urbanization and the collection of surpluses are indicative of what the sociologist Mann (1984) calls increased “infrastructural power” of the state: an authority that imposes itself through bottom-up action in the daily life of the inhabitants.

The control of burial rites and “expenditure” suggests that political power at Pontecagnano predated the appearance of the city as centralized settlement proper. An authority was invested with the power to control burial and thus collective memory. The authority might have had a certain control over resources and redistribution of resources, such as bronze, as well. Even though scoring “low” on Michael Mann’s scale of despotic and infrastructural power (Mann, 1984), this type of power is characteristic for “the state.”

The early state at Pontecagnano, moreover, must also have had a say on the use of space, given the close proximity of the burial plots. Communal space, e.g., grazing grounds, quarries for clay, and stone as well as field boundaries had to be managed. This means that, despite the physical fragmentation of settlement nuclei and burial plots, “the state” at Pontecagnano was spatially bound and consisted of interconnected nuclei—i.e., a network.

We often think of settlement centralization as the moment of “birth” of the city. This is also reflected in the nomenclature “pre-urban,” “proto-urban,” and “urban.” However, as the present analysis suggests, change at Pontecagnano is not radical and...
clearly part of a broader process of state transformation rather than formation. Settlement centralization is just one strategy adopted by the state to exercise infrastructural power. It just happens to be a transformation that has a high archaeological visibility and thus appears to constitute a radical break with the past. Alternatively, studying different settlement nuclei as part of a network in perpetual transformation allows to capture its social functioning, interdependence and manipulation of space much better. One can, therefore, wonder if the term “low density urbanism” might not be more appropriate to describe the situation in Early Iron Age Italy, rather than the terms “pre-urban” and “urban,” which are a priori invested with specific values. “Low density urbanism” has been investigated for prehistoric settlement processes in Central Europe as well as the Maya in Mesoamerica and Iron Age oppida (Fletcher, 2011; Chapman and Gaydarska, 2016; Moore, 2017) and it places emphasis on the connectedness of settlement within a social system, rather than the categorical value of space and settlement density. “Low density urbanism” thus provides a more productive anchor point for comparative analysis and interdisciplinary dialogue on the formation of the Ancient City.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This paper was published thanks to OPEN-AIRE funding granted to Francesca Fulminante for the Marie Sklodowska IEF Past-People-Nets 628818 conducted at the University of Roma Tre (2014–2016).

ACKNOWLEDGMENTS

I wish to thank the funding bodies that have made my research during the last years possible: Georg-August-Universität Göttingen (Dorothea Schlözer program 2014–2016) and the Horizon 2020 program (Marie Sklodowska Curie Fellowship grant no. 702511 in 2016–2018). My thanks also goes to the peer reviewers who helpedfully commented on an earlier version of this text.

REFERENCES


Blake, E. (2014). Social Networks and Regional Identity in Bronze Age Italy. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139897926


doi: 10.1037/10648-000


Copyright © 2019 Donnellan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
A New Kind of Relevance for Archaeology

Scott G. Ortman *

Department of Anthropology, University of Colorado Boulder, Boulder, CO, United States

Traditional ways of doing archaeology impact the world in a variety of ways, but despite recent efforts the practical relevance of archaeology has remained limited. In this paper, I discuss why this is the case and suggest how archaeology can achieve greater practical relevance. I argue, first, that the traditional focus of archaeology on reconstructing the past is valuable but is unlikely to expand its practical relevance because the results are too context-specific. Second, I suggest traditional responses to the problem of historical contingency are also inadequate because the results are too general to connect to the specific issues and solutions society needs. Finally, I make a surprising and perhaps radical suggestion: that a productive way forward is to resuscitate and reformulate elements of the New Archaeology that were never realized by its proponents. I use the example of settlement scaling theory to illustrate that this is both possible and productive, and that additional work in this spirit would enhance the practical relevance of our field.

Keywords: archaeological synthesis, contemporary relevance, settlement scaling theory, agglomeration, archaeological theory and method

INTRODUCTION

In recent years an increasing number of archaeologists have conducted research that is explicitly designed to address contemporary issues (Sabloff, 2008; Cooper and Sheets, 2012; Ingram and Gilpin, 2015; Nelson et al., 2015; Liebmann et al., 2016; Kohler et al., 2017; Hambrecht et al., 2018; Hegmon and Peeples, 2018; Jackson et al., 2018). Despite many exciting results emanating from this work, as of yet it seems to have had little impact on actual public policy discussions. For example, despite extensive research by archaeologists on human responses to climate change, to date the results of such research have been largely absent from reports by the Intergovernmental Panel on Climate Change (Jackson et al., 2018). Given that the archaeological record is the most extensive compendium of human experience there is, it seems only natural that the results of archaeological research should have an impact on discussions concerning contemporary issues (Smith et al., 2012; Kintigh et al., 2014; Altschul, 2016; Altschul et al., 2017). But so far there seems to have been limited success in this regard. Why is this? What would an archaeology that has practical relevance beyond archaeology look like? How would it be different from the archaeology many of us practice right now?

In this essay I will offer my own opinions on these sorts of questions. I will argue, first, that the traditional focus of archaeology—constructing historical narratives—is valuable but is unlikely to expand its practical relevance because the results are too contingent on local details. Second, I will argue that traditional “grand synthesis” and cross-cultural research are also insufficient because their results are too general to connect to specific issues and solutions. Finally, I will suggest, perhaps surprisingly, that a productive way forward is to resuscitate and reformulate aspects of the New Archaeology that were not realized in the 1970s. I use the example of settlement scaling theory
to illustrate that the New Archaeology’s interest in developing predictive knowledge of specific social phenomena is both possible and productive, and that additional work in this spirit may be our best way forward. In a nutshell I believe that, if archaeology is to achieve greater practical relevance it will not be through research that reconstructs the past or makes broad generalizations. Rather, it will come from studies of specific social phenomena regardless of where or when they occur.

WHAT IS PRACTICAL RELEVANCE?

Before getting into the main arguments of this paper, I should say a few words about what I mean by practical relevance. There are many aspects of archaeology that yield practical benefits in the present, from developing sites for cultural tourism to creating the raw material for museum exhibits to promoting social justice for marginalized groups. Here, I use the term “practical relevance” to refer to something more specific: predictive knowledge of specific social phenomena that can help us make informed decisions regarding issues we face today.

Two questions come immediately to mind. First, is it really worthwhile to view human behavior as predictable? There are of course many aspects of the behavior of individuals, and of groups, that are not predictable. But at least some are. As examples: today’s demographic rates have predictable effects for tomorrow’s economy; insurance companies use actuarial tables to predict payouts and adjust premiums with reasonable confidence; political scientists create models based on demographic and socioeconomic characteristics of subgroups that predict election results; the daily movements of individuals follow predictable patterns that allow our smartphones to plot the most time-efficient route of travel between two places; simple models often surpass expert judgment in predicting the outcomes of sporting events; and tech companies use browsing and posting habits to predict which products we are most likely to purchase.

It’s also important to recognize that the ability to predict is generally beneficial. Knowing how many people of different ages will be around at a future date is critical for maintaining the finances of the social safety net; actuarial tables ensure that insurance companies can honor their commitments to people in need; predicting travel times helps individuals use their time more effectively, and connecting people with the products they are likely to want helps consumers in addition to tech companies. So even though many aspects of human behavior may never be entirely predictable, at least some are, at least partly, and it therefore stands to reason that social scientists should be able to expand knowledge of predictable behavior with appropriate effort.

Second, even if one grants that human behavior is at least partly predictable, is it really reasonable to imagine that the knowledge generated through archaeology is relevant for issues we face today? After all, societies of the past were different in innumerable ways from those of the present. They were smaller, lacked modern transport and information technologies, had different social and political institutions, and operated in terms of diverse cultural concepts that for the most part do not characterize late-stage capitalist nations of the present. Given all these differences, why should anyone think the results of archaeology actually apply to today’s decisions?

One possible answer involves social theory. For many decades archaeologists have engaged with social theorists in cultural anthropology, sociology, geography, and related fields to make ontological claims regarding sociocultural phenomena, and in many cases these frameworks have been devised in the context of contemporary societies and then applied to past societies by archaeologists (Shanks and Tilley, 1987; Trigger, 1989; Hodder, 1991, 2012; Olsen, 2010; Alt and Pauketat, 2019). So there is an established tradition that argues, in effect, that the basic properties and relations of human social life apply to all societies. This approach has yielded many insights, but it seems limited from the perspective of practical relevance in that the approach generally does not lead to predictions that can be evaluated empirically. Instead, in most cases the process involves mapping or indexing a conceptual framework onto archaeological information from a given context (Smith, 2015, 2017). Most of the time, this approach helps one interpret the archaeological evidence better, but it doesn’t lead to empirical predictions such that one can know if or how a particular idea is wrong.

By and large, the social sciences do not yet possess a body of such ideas, and I suspect many archaeologists would question whether it is even possible. We should be under no illusions that developing a predictive theory of human society is easy. Still, the history of other sciences provides a basis for optimism. Newtonian mechanics applies to all objects and has sufficient predictive power to engineer spaceships that get people to the moon and rovers to Mars. The periodic table applies to all elements and makes it feasible to develop new compounds. The Neo-Darwinian synthesis leads to predictions about how populations of organisms as simple as bacteria, and as complex as human beings, change from one generation to the next. And some would even argue that behavioral economics reflects intrinsic aspects of human cognition and leads to predictions about human decision-making in any context (Kahneman, 2011; Thaler, 2016). Developing these frameworks is hard—the very fact that the scientists most responsible for these insights are household names is a hint of the difficulties involved. But we need to do it, and more importantly, we need to believe it is a good thing to do, if the social sciences are to play a more important role in our future.

With this perspective in mind, what would an archaeology that has practical relevance for today look like? Recognizing that ultimately this will require theoretical development, I focus in the following sections on the epistemological and methodological basis of what I believe would be a productive approach. I will approach this vision by first illustrating why several traditional approaches to archaeological interpretation are unlikely to achieve practical relevance. Then, I’ll suggest that archaeology got close to moving in this direction in the 1970s but for various reasons turned away from it. Finally, I’ll develop an example which illustrates that practical relevance can be achieved if we are willing to apply the same reasoning and analytical processes that are used throughout the sciences to the
material proxies for human behavior we can now derive from the archaeological record.

THE RELEVANCE OF HISTORY

Archaeologists are good at historical reconstruction, and getting better all the time. I’d like to think I’ve contributed to this effort myself (Ortman, 2012, 2016a; Ortman and McNeil, 2018). From GIS to AMS-dating, isotopes, ICP-MS, micromorphology, phytoliths, ancient DNA, LiDAR, UAV photogrammetry, linguistic paleontology and more, archaeological methods continue to expand our ability to reconstruct past human behavior and environments; and as these methods expand, our historical narratives become increasingly detailed, accurate, and compelling. Today, we really do know far more about the details of the human experience through archaeology than we did a few decades ago. And these narratives are important. They feed our imaginations regarding the range of social worlds humans have created and the range of worlds that are possible (e.g., Fowles, 2013); they provide both celebratory and cautionary tales regarding what can happen (Harper, 2017); they support human diversity and multi-culturalism by illuminating the heritage of contemporary peoples (Popa, 2019); and they even contribute to social justice by getting the facts of history right (Preucel, 2002; Cameron, 2008). All of these outcomes are valuable and I want to stress that, in making the argument of this paper, I do not mean to suggest that archaeologists should not continue doing good work in all these areas. Instead, the question I wish to ask is whether the historical narratives that most archaeologists contribute to can lead to predictive knowledge that might help us make informed practical decisions: how (or if) to define land-use zones, redistribute wealth, stimulate economic growth, reduce environmental impacts, improve public health, mitigate the effects of climate change, and so forth. In other words, my question is whether such narratives give us a basis for predicting future outcomes based on actions we could take today.

An example may help to illustrate what I have in mind. Let’s say that one is interested in using the results of archaeological research to suggest productive ways of adapting to climate change. One way to proceed would be to examine specific cases where the long-term history of human-climate relations is understood in great detail. A good example is The Village Ecodynamics Project (VEP), an interdisciplinary project involving archaeologists, computer scientists and ecologists that has worked since 2003 to examine human-environment relationships in the US Southwest. The project has received substantial financial support from agencies and organizations in the US, and I have been fortunate to be a part of it. Working together, we have retrodicted past precipitation and temperature in two study areas by correlating tree-ring series with weather station and pollen core data (Wright, 2012; Bocinsky and Kohler, 2014), and then then translated these into productivity estimates (at a temporal resolution of 1 year and a spatial resolution of 4 ha) by combining paleoclimate reconstructions with soils and historic crop yield data (Kohler et al., 2007, 2012b; Varien et al., 2007; Bocinsky and Varien, 2017). We also compiled architectural, ceramic, and chronometric data for thousands of archaeological sites and used these data to estimate the population histories of our study areas at a temporal resolution approaching a single human generation (Ortman et al., 2007; Ortman, 2016b; Schwindt et al., 2016). We created time series for interpersonal violence rates, demographic rates and hunting pressure on wild game (Johnson et al., 2005; Kohler et al., 2008, 2009, 2014; Kohler and Reese, 2014), and we reconstructed patterns of settlement, community organization and migration into and out of our study areas (Glowacki and Ortman, 2012; Ortman, 2012; Glowacki, 2015; Kemp et al., 2017). Finally, we developed agent-based models that provide robust null models for assessing the effects of climate for demographic rates and social organization (Kohler, 2012; Kohler et al., 2012a; Crabtree et al., 2017).

Through this work we have developed an incredibly-detailed reconstruction of the social and environmental history of the ancestral Pueblo people who lived in our study areas. Indeed, I think it is fair to say that our syntheses of the archaeological and environmental records of these areas incorporate more cumulative expenditures on archaeology than for any comparably-sized areas anywhere in the world. As a result, we now have a much clearer picture regarding how this society collapsed around 1280 CE. Several centuries of rapid population growth, in the context of a subsistence farming society with a modest division of labor, led to a substantial fraction of the population living on land that was vulnerable to drought. When drought finally hit, the overall landscape was still productive enough to feed the population, but people who lived on the most productive lands were not accustomed to producing food surpluses, and people who needed food the most had no means of obtaining it through the economic system. Social breakdown, characterized by extreme internece violence and a rejection of existing social institutions, led to mass migration and the end of a cultural tradition.

Based on VEP research, it is now clear that the social response to drought was far in excess of its actual impact to regional agricultural potential. And the organization of the society seems to have been a primary reason. Indeed, it is tempting to conclude from this work that a good way to ameliorate the social consequences of climate change is to promote development of non-agricultural sectors in developing nations. But here is where the problem with history begins to show itself. There are competing views on just about every issue in society, even among those who are committed to fact-based analysis. So it is not difficult to imagine someone cross-examining the VEP research by asking “How do you know from this specific case that there is a predictable relationship between climate change, level of economic integration, and extent of sociopolitical disruption?” At this stage, the only honest response would be “we don’t.” Despite all our efforts to get the details of history right in this case, and the exceptional investment of resources in doing so, in the end we cannot say whether the observed level of sociopolitical disruption is a predictable outcome of general processes or a contingent outcome of specific circumstances. We hoped our agent-based models might do this, and they do seem to account for certain aspects of this history, but none of these models
reproduce the most obvious and important outcome, which is the actual collapse of the society.

**TRADITIONAL RESPONSES**

This is just one example, but I think it serves to illustrate the point that historical reconstructions always arrive at the same destination. When history is the goal, increasing research time and effort inevitably lead to greater focus on local details at increasing levels of magnification. We do learn a lot more about the operation of practical reason, while others reflect little-understood tendencies of the human mind to produce particular economic than strictly ecological or cultural (Trigger, 2003, p. 674–676). “Some of the parallels appear to result from the operation of practical reason, while others reflect little-understood tendencies of the human mind to produce particular types of analogies” (Trigger, 2003, p. 685). These are deep insights, but they are very general, and as such they do not provide much basis for practical decisions one could make to address a specific contemporary issue. So although traditional synthesis yields fascinating generalizations, it is not structured enough to provide more than a starting point for an archaeology with practical relevance.

The second approach is cross-cultural analysis. As with traditional synthesis, there is a long and varied tradition in this sort of work, in both cultural anthropology and archaeology (Childe, 1936; Adams, 1966; Ford, 1969; Blanton et al., 1993; Johnson and Earle, 2000; Trigger, 2003; Diamond, 2005; Flannery and Marcus, 2012; Jennings, 2016). Such studies have always identified interesting patterns, at least some of which must reflect predictable regularities in human affairs. But due to the inter-correlated nature of many properties of human societies it remains extremely difficult to identify predictable causal pathways that relate to specific issues. To offer just one example: in *Understanding Early Civilizations*, Trigger (2003) found that early civilizations exhibit idiosyncratic cultural variation but strong regularities in their economies and social and political organizations that cannot be explained by historical connections or shared ancestry. He concludes that the primary factors behind the emergence of civilization are more political and economic than strictly ecological or cultural (Trigger, 2003, p. 674–676). “Some of the parallels appear to result from the operation of practical reason, while others reflect little-understood tendencies of the human mind to produce particular types of analogies” (Trigger, 2003, p. 685). These are deep insights, but they are very general, and as such they do not provide much basis for practical decisions one could make to address a specific contemporary issue. So although traditional synthesis yields fascinating generalizations, it is not structured enough to provide more than a starting point for an archaeology with practical relevance.

The second approach is cross-cultural analysis. As with traditional synthesis, there is a long and varied tradition in this sort of work, in both cultural anthropology and archaeology (Murdock, 1949; Driver and Massey, 1957; Oliver, 1962; Carneiro, 1967; Jorgensen, 1980; Ember and Ember, 1994; Peregrine, 2003; Gell-Mann, 2011; Whitehouse et al., 2019). Much of it has involved extraction of nominal or ordinal variables from primary ethnographic and archaeological literature that was rarely created of written for this purpose. For the most part, these studies focus on establishing statistical relationships between variables. A good recent example of this style of research is the SESHAT project, which has compiled a global archaeological and historical database and used it to test hypotheses about the underlying structure of variation in human social organization at the level of polities (Turchin et al., 2018). SESHAT researchers collected data for 51 (nominal, ordinal and continuous) variables from 414 polities dating from 9600 BCE to 1900 CE and aggregated these into nine “complexity characteristics”: polity population, polity territory, capital population, hierarchy, “texts,” information system, infrastructure, money, and government. Principal components analysis of the scores for each of these characteristics shows that all are highly correlated, indicating that they all tend to evolve together.

This is a strong finding that expands knowledge of the general process of human social evolution. More importantly for the purposes of this paper, the results allow one to predict that if one dimension of social complexity increases, the others are more likely than not to follow suit. Still, notice that what is being predicted in this case is a correlated increase in measures that are complex combinations of many nominal, ordinal, and/or continuous variables. As a result, from this analysis it is not possible to determine how a certain amount of change in any specific property will affect any other property. This is what would be needed for these results to have practical value in addition to scientific value. Also, since the unit of analysis is the polity, many problems related to the internal functioning of societies, cities or households cannot be addressed. So although cross-cultural analysis can lead to predictive knowledge, such studies tend to operate at a level of abstraction that is too general to address specific social problems and solutions.

These two traditional responses to the problem of historical contingency, then, have the opposite problem: instead of leading to results that are too contingent on local and historical factors to apply elsewhere, they lead to results that are too general to be useful for predicting the outcomes of specific actions. Identifying cross-cultural regularities and patterns in (pre)history is extremely interesting, and one would expect much useful information to be embedded in the results of such studies. But the relationships identified through such studies are typically too general for practical application.

**UNFINISHED BUSINESS?**

What then to do? I’d begin by noting that the issues discussed above are once again nothing new, as archaeologists have been aware of the shortcomings of traditional approaches as a means of generating predictive knowledge of human affairs ever since the foundational writings of the New Archaeology. This intellectual movement of the 1960s and 70s drew on the philosophy of logical positivism, which was viewed by its proponents as the foundation of the natural sciences, in an attempt to generate “covering laws” that applied to the entirety of the archaeological record (Hempel, 1966; Binford, 1968; Watson et al., 1971). The New Archaeology was not successful in its stated aims, but I want to suggest that the reasons behind its failure may help archaeology chart a path toward enhanced practical relevance.

The New Archaeology had several shortcomings. One was the appeal to philosophers of science as opposed to actual
practice. This was unfortunate because logical positivism is an abstraction that never characterized actual practice in the natural sciences (Smith, 2017). To give just one example, contrary to the formal, binary logic of logical positivism (“In C, if A, then B") (Watson et al., 1971, p. 6–7), most scientific knowledge claims are actually statistical: what the average outcome should be, the likelihood of a certain level of effect, and so forth. A second shortcoming was a faulty conception of “explanation.” In its best-known manifesto, Watson et al. (1971) argued that the major goal of archaeology was to show that specific past events are instances of a general or “covering” law. In their words, “A scientist explains a particular event by subsuming its description under the appropriate confirmed general law, that is, by finding a general law that covers the particular event by describing the general circumstances, objects, and behavior of which the particular case is an example” (Watson et al., 1971, p. 5). This formulation suggests the goal of archaeology is to explain the specific historical case by showing that it is an instance of a general rule. Archaeologists can certainly do this. But the earlier discussion of history suggests that if the goal of archaeology is to explain the specific event, delving into the details toward historically-contingent factors will be far more productive. So following this procedure actually drives one away from the search for generalizations that have practical relevance.

A final shortcoming of the New Archaeology was a fuzzy distinction between explanation of human behavior vs. explanation of the archaeological record (Schiffer, 1972). Explaining why the archaeological record has the properties it has—what has come to be known as middle-range theory—is a necessary step in translating observations of that record into proxies for past human behavior. But such theory was largely absent in the 1960s, and as a result early attempts at explanation in archaeology, notably the work of the so-called “ceramic sociologists” (Longacre, 1964; Hill, 1970), were readily deconstructed (e.g., Allen and Richardson, 1971). Still, several aspects of contemporary archaeology, including the study of site formation processes, taphonomy and ethnoarchaeology, are positive outcomes. In the US Southwest, for example, archaeologists today routinely use generalizations derived from ethnoarchaeological studies of abandoned structures to interpret the fill stratigraphy and floor assemblages of ancient dwellings (Stevenson, 1982; Schiffer, 1985; Cameron and Tomka, 1993); and they use the discard equation to relate artifact accumulations to household inventories, people, and time (Schiffer, 1987; Mills, 1989; Varien and Mills, 1997; Varien and Potter, 1997; Varien and Ortman, 2005). The relationships between human behavior and site formation processes captured in these approaches are highly predictable; indeed, one can rightly claim that these relationships explain basic properties of the archaeological record.

But in the end, explaining the archaeological record as a present-day phenomenon is only an instrumental goal in and of itself. It’s a necessary step, but things only start to get relevant outside of archaeology when one uses this knowledge to study human social dynamics. Since none of this existed in the 1960s, proponents of the New Archaeology quickly realized that middle-range theory had to come first, and as a result the scientific knowledge they produced focused on the archaeological record as a present-day phenomenon. Kent Flannery famously derided the initial results as mere “Mickey Mouse laws” (Flannery, 1973), and such critiques led archaeologists to abandon the ultimate goal of the New Archaeology program and return to the goals of traditional synthesis, leading (among other things) to the variety of evolutionary approaches that continue to have practitioners (and critics) today (Wright and Johnson, 1975; Sanders et al., 1979; Flannery and Marcus, 1983, 2012; Feinman and Marcus, 1998; Johnson and Earle, 2000; Laland and Brown, 2002; Shennan, 2002; Smith, 2003; Yoffee, 2005; Pauketat, 2007; Jennings, 2016; Lekson, 2018).

Much of this intellectual history is well-known to archaeologists, and the field has advanced in many ways since the 1970s. Still, notice what the ultimate goal of the New Archaeology actually was: to discover regularities in human social behavior that are context independent, with the implication that they apply to the present as well as the past. And notice what its methodology was: to develop theory that leads to predictions (“test expectations”) that can be checked against measurements derived from the archaeological record. This sounds precisely like the kind of knowledge that would contribute to contemporary conversations regarding urban planning, economic development, inequality, sustainability, migration, health, and other issues. In short, the New Archaeology would appear to represent an initial, and still unrealized, attempt to achieve practical relevance for archaeology. In the process of thinking through what archaeology as a social science would look like, it became apparent that archaeologists needed to translate material traces into reliable proxies for past human behavior before it could hope to investigate human social dynamics. Archaeologists today routinely use the results of this effort as part of normal practice. Perhaps the issue, then, is that it was not possible to realize the ultimate goal of the New Archaeology because the field needed to develop middle-range theory first. In other words, perhaps the failure of the New Archaeology was not due to a mismatch between scientific reasoning and human society; but because archaeology had to build the capacity to study human social dynamics before it could apply such reasoning to the study of specific social phenomena. The New Archaeology was highly successful with this initial goal. Perhaps we can still accomplish the second?

I suspect many readers will have an immediate negative reaction to this suggestion. The New Archaeology was clearly not successful in its stated aims. And I suspect many readers would argue that the reason it failed is because a natural science-type of reasoning does not apply to human affairs. After all, archaeology is a historical science, like paleontology, where it’s not possible to achieve experimental control or re-run the tape of history again and again (Gould, 1989). And the archaeological record is hopelessly haphazard and partial in its details. The material residues of past behavior that it preserves vary dramatically for all manner of reasons, from the material cultures and technologies of past societies to subsequent disturbance to decomposition and so forth. It’s also quite expensive to collect enough data, in systematic enough ways, to really use this record in a natural science kind of way. So we shouldn’t pretend we can. And so the argument goes.
But let’s think about this argument a bit more, using an example of how scientific research is actually practiced in a field that generates useful knowledge. Although it does not provide a perfect analogy, the example of clinical trials is instructive. When medical researchers test the efficacy of a new drug, they typically study three groups—one that receives the treatment, a second that receives a placebo and a third that does not receive a treatment at all. As the patients are human beings with free will, it is impossible to completely control for variation in human biology, the life history of patients prior to treatment, and the behavior of patients during or after treatment. So, in clinical trials “experimental control” is achieved by stratifying patients into genetic, demographic and/or lifestyle subgroups and then examining the effect of the treatment across large numbers of people in each group, under the assumption that the uncontrolled effects will effectively cancel each other out. There is no attempt to quantitatively or even document all of these uncontrollable factors.

In other words, variation in the biology and experience of individuals is at best only partly controlled in such experiments. Instead, experimental control is achieved through sample size and stratification into subgroups. The logic of such studies is that despite the myriad uncontrollable factors that govern outcomes for any given individual, it is still possible to determine the average effect of a single factor across a population, and to determine courses of action that have a significant impact on peoples’ lives, through statistical analysis of outcomes for many individuals across subgroups. The main methodological principle in clinical trials, then, is that to learn something useful about a particular unit of study (in this case, individual humans), the best way to control for all the factors that one cannot control at the level of that unit of study is to compare results across large numbers of units. When this is done, one can develop predictive knowledge concerning the average effects of a specific factor for specific outcomes. And the results are clearly useful. Indeed, in the case of clinical trials, many peoples’ lives depend on them.

Notice that the practical relevance of clinical trials does not necessarily derive from exotic analytical or statistical methods. Indeed, the statistical techniques typically used in clinical trials (statistical tests, regression, etc.) are also part of the basic toolkit of archaeologists. And using these tools, it is possible to say that we “know” that a certain type of pottery is older than another; that the average house grew larger over time; or that the length of a knife is unrelated to its width. In other words, this logic, which characterizes both clinical trials and archaeology, can and does lead to secure and even predictive knowledge of the world. For example, we can use it to predict, with high confidence, that an archaeological site at which a certain variety of pottery is common was occupied during a certain time period.

The main difference between artifact analysis and the clinical trial, then, is the practical relevance of the unit of analysis. Knowing that a certain treatment will increase life expectancy for patients makes a difference in peoples’ lives today. Knowing that sites bearing a certain kind of pottery were inhabited during a certain period, in and of itself, does not. The point here is that archaeologists know how to do this kind of analysis; we just don’t typically do it in such a way that the results could have practical relevance. For archaeologists, our potentially relevant units include households, neighborhoods, settlements, polities, ethnic groups, and populations. But in most studies of these units, we have been content with historical reconstruction, traditional synthesis, or cross-cultural comparison. We do not have a tradition of applying the same techniques we normally apply in everyday analysis and interpretation to the units that matter beyond our field.

What I am suggesting, then, is that what archaeology needs to do to achieve greater practical relevance is replace the patient in the clinical trial with a household, neighborhood, settlement, polity, ethnic group, or population, based on relevant material proxies supported by middle-range theory. There is no reason why archaeologists cannot do this. We just need to apply this logic to relevant units of analysis, design and implement appropriate methodologies, and use the law of large numbers to provide effective controls. In addition, we need to work with other social scientists to develop theories, models, and expectations regarding how proxies for human behavior derived from the archaeological record might be expected to vary under specific conditions. There is no road map for doing this, but in the final section I’d like to develop an example, drawn from my more recent work, which shows that this can be done.

**AN EXAMPLE**

In this final section, I discuss the ideas, methodology, and results of the Social Reactors Project, a collaboration among archaeologists, urban scientists, and economists that is investigating agglomeration effects, past and present, using ideas from network science and complex systems. I present this example not so much to promote these particular ideas (although I do find them compelling), but to illustrate the more general point that it is possible to develop predictive knowledge of human affairs that is relevant for the present and future by combining familiar archaeological proxies and analytical methods with a dose of theoretical abstraction.

The basic idea at the center of our approach is that when humans arrange themselves in space, they do so in ways that balance the material benefits of social contact with the cost of moving around to do it. We do not view this as a utility maximizing process (as in economics), but as a balancing of costs and benefits following the tradition in geography. We suggest the spatial equilibrium resulting from this balancing act leads to the concentration of humans, their interactions, and their outcomes, in space and time. As a result, individuals in larger settlements have more social contacts and exchanges per unit time; and there are also increased opportunities for specialization as individuals can meet more of their material needs through human networks as opposed to their own individual effort. This process, which we label the “social reactor process,” induces human networks to grow in consistent, non-linear, and open-ended ways with population.

The main difference between artifact analysis and the clinical trial, then, is the practical relevance of the unit of analysis. Knowing that a certain treatment will increase life expectancy for patients makes a difference in peoples’ lives today. Knowing that sites bearing a certain kind of pottery were inhabited during a certain period, in and of itself, does not. The point here is that archaeologists know how to do this kind of analysis; we just don’t typically do it in such a way that the results could have practical relevance. For archaeologists, our potentially relevant units include households, neighborhoods, settlements, polities, ethnic groups, and populations. But in most studies of these units, we have been content with historical reconstruction, traditional synthesis, or cross-cultural comparison. We do not have a tradition of applying the same techniques we normally apply in everyday analysis and interpretation to the units that matter beyond our field.

What I am suggesting, then, is that what archaeology needs to do to achieve greater practical relevance is replace the patient in the clinical trial with a household, neighborhood, settlement, polity, ethnic group, or population, based on relevant material proxies supported by middle-range theory. There is no reason why archaeologists cannot do this. We just need to apply this logic to relevant units of analysis, design and implement appropriate methodologies, and use the law of large numbers to provide effective controls. In addition, we need to work with other social scientists to develop theories, models, and expectations regarding how proxies for human behavior derived from the archaeological record might be expected to vary under specific conditions. There is no road map for doing this, but in the final section I’d like to develop an example, drawn from my more recent work, which shows that this can be done.

**AN EXAMPLE**

In this final section, I discuss the ideas, methodology, and results of the Social Reactors Project, a collaboration among archaeologists, urban scientists, and economists that is investigating agglomeration effects, past and present, using ideas from network science and complex systems. I present this example not so much to promote these particular ideas (although I do find them compelling), but to illustrate the more general point that it is possible to develop predictive knowledge of human affairs that is relevant for the present and future by combining familiar archaeological proxies and analytical methods with a dose of theoretical abstraction.

The basic idea at the center of our approach is that when humans arrange themselves in space, they do so in ways that balance the material benefits of social contact with the cost of moving around to do it. We do not view this as a utility maximizing process (as in economics), but as a balancing of costs and benefits following the tradition in geography. We suggest the spatial equilibrium resulting from this balancing act leads to the concentration of humans, their interactions, and their outcomes, in space and time. As a result, individuals in larger settlements have more social contacts and exchanges per unit time; and there are also increased opportunities for specialization as individuals can meet more of their material needs through human networks as opposed to their own individual effort. This process, which we label the “social reactor process,” induces human networks to grow in consistent, non-linear, and open-ended ways with population.

The basic idea at the center of our approach is that when humans arrange themselves in space, they do so in ways that balance the material benefits of social contact with the cost of moving around to do it. We do not view this as a utility maximizing process (as in economics), but as a balancing of costs and benefits following the tradition in geography. We suggest the spatial equilibrium resulting from this balancing act leads to the concentration of humans, their interactions, and their outcomes, in space and time. As a result, individuals in larger settlements have more social contacts and exchanges per unit time; and there are also increased opportunities for specialization as individuals can meet more of their material needs through human networks as opposed to their own individual effort. This process, which we label the “social reactor process,” induces human networks to grow in consistent, non-linear, and open-ended ways with population.
The key question, for the purposes of this essay, is how we justify the claim that the social reactor process is an intrinsic property of human settlements. After all, there are innumerable social, cultural, geographic and historical factors, beyond population, which interact in complex and often unobservable ways to produce the observable properties of each individual settlement. How can one claim to know that population, by itself, has a predictable effect on such properties? There are two parts to the answer. First, we use the results of middle-range research to identify archaeological proxies for the parameters of settlement scaling models. The sorts of measures we have used include house and structure counts and densities; the lengths and widths of roads, paths and public spaces; the areas and volumes of houses and public works; and the densities, ratios, and diversity of artifact types.

Second, we use the logic of the clinical trial. The archaeological record is obviously haphazard when viewed in detail. Not only is preservation partial, but investigation of the remaining traces is also biased in several ways due to the time and expense involved in archaeological field and laboratory work and the changing interests of investigators over time. As a result, there is error associated with every measurement, and we cannot know, for example, the exact momentary population, or the precise rate of pottery consumption, for any past settlement. And more importantly, even if we could measure the properties of individual settlements precisely and accurately, it would still be the case that every settlement has a unique history, such that a myriad of factors beyond population, only some of which are observable, have combined in unique ways to produce its specific observed properties. Due to these combined effects of measurement error and historical contingency, it is not reasonable or feasible to test predictions of settlement scaling theory (SST) through analysis of a single settlement. The only way to do it is to compare many settlements, ideally from many settlement systems, to see if the predicted effects are apparent, on average, across all of them.

It turns out this task is relatively straightforward once one has compiled relevant data. SST argues that the average effect of settlement population for an aggregate property of interest is given by a power function \( Y = Y_0N^\beta \), where \( Y \) is the aggregate property, \( Y_0 \) is a baseline value, \( N \) is the settlement population, and \( \beta \) is an exponent that summarizes the rate of increase of the property relative to the population. The theory also includes mathematical models that derive predictions for what the exponent \( \beta \) should be, depending on whether the property of interest represents a socio-economic rate, a measure of functional diversity, or a measure of physical infrastructure. These scalar effects of human networks can be observed empirically by fitting a linear function to log-transformed measures of \( N \) and \( Y \) across a sample of settlements in a system. This is feasible because \( Y = Y_0N^\beta \) and \( \log Y = \beta \log N + \log Y_0 \) are equivalent expressions. When this is done, the slope of the fit line is an estimate of \( \beta \), and its intercept is an estimate of \( \log Y_0 \), and thus of \( Y_0 \). The details of SST have been presented in a variety of places (Bettencourt, 2013, 2014; Ortman et al., 2014, 2015; Youn et al., 2016), but for present purposes the key point is that the analysis determines whether, on average, the estimated exponent \( \beta \) falls within the range of statistical tolerance of the value predicted by the relevant model. So when we conduct a scaling analysis, we are testing whether a specific prediction of the framework is borne out by the data. When the data do not conform to the prediction, it tells us that something is wrong, either with the model or with the data.

An example of such a test is shown in Figure 1, which examines the relationship between settlement population and aggregate settlement productivity in the archaeological record of five New World societies: the Basin of Mexico; the Prehispanic Upper Mantaro Valley of highland Peru; the Mesa Verde region in Colorado, USA; the Middle Missouri region in North and South Dakota, USA; and the Lower Santa valley of coastal Peru (Data Sheet 1). The data for the Lower Santa valley derive from a settlement pattern survey of the region by Wilson (1985, 1988), and the other datasets have been analyzed in previous publications (Ortman et al., 2015, 2016; Ortman and Coffey, 2017). These settlements encompass six orders of magnitude in population, \( 60^2 \) in latitude, and 6,000 years in time.

The proxy for settlement population varies across societies. In the Basin of Mexico population is estimated either by multiplying the domestic mound count by the average household population, or by multiplying the site area by a population density indexed to the surface artifact density; in the Middle Missouri, Upper Mantaro and Lower Santa the estimated population is simply the number of domestic residences in the settlement; and in the central Mesa Verde the estimated population is the number of pit structures present. In all cases, the proxy for a socio-economic rate is the total area of the domestic structures (or mounds) in the settlement. We treat the latter as a measure of total settlement productivity based on a variety of archaeological and ethnoarchaeological studies which support an association between house size and wealth (Smith, 1987; Blanton, 1994; Kohler and Smith, 2018). The basic argument is that, because most wealth in past societies took the form of tangible goods, households that had more stuff per person needed more floor area per person. A recent demonstration of this comes from a study of households in Aztec-period Central Mexico which found that larger houses are associated with greater amounts of more valuable possessions (Olson and Smith, 2016).

The process by which total house areas are estimated varies substantially across settlements both within and between regions. In some cases, total roofed space was measured directly based on complete surface preservation or geophysical survey; in others counts and average areas of different classes of structure are reported; and in still others only the counts and areas of those domestic mounds that happen to be preserved are reported. In such cases we either multiplied the average mound area by the house count to estimate the total space, or we calculated a weighted average area per structure based on reported counts and average areas of documented structure types and then multiplied the total structure estimate by this average.

Due to the realities of the archaeological record, and the resulting data, there are obviously errors in the estimates of
FIGURE 1 | The relationship between settlement population and total house area in five New World societies. Note that the intercept of the fit line for each society is different, but the slope (coefficient) of the fit line is very similar across cases.

Table 1 shows that these slopes are all in excess of one and in the vicinity of the theoretical prediction of 7/6, or 1.167. All of the regressions have high $r^2$-squared values, but these are in part autocorrelation effects that derive from using the house count to construct the roofed space estimate at many sites. Still, in most cases the 95% confidence interval of the estimate for beta excludes one, which is what the slope of the relationship would be if the estimates of roofed space per capita were independent of the site population. These results thus provide striking evidence for a specific empirical regularity in the relationship between population and material productivity.

This uniformity can be made even clearer by centering the data from each region so that the mean coordinate of each dataset is at the origin. This is done using the following formula:

$$\text{centered } (x_i) = x_i - \left( \frac{\sum_{i=1}^{n} x_i}{n} \right),$$

which allows one to use the data from all five regions in a single regression analysis. The relationship for the centered data is presented in Figure 2, and in the bottom row of Table 1. This analysis leads to a remarkable result. The value of $\beta$ predicted
Evidence that house areas follow a single scaling relationship across societies. In this plot, the data have been centered by subtracting the mean coordinates of the data for each society from each data point. This process re-scales the data so that their center of the data for each society is at the origin. The estimated coefficient of the scaling relationship for the centered data is within two one thousandths of the predicted value.

by SST for socio-economic rates is 7/6, or 1.167; the observed value in this centered dataset is 1.165, with a standard error of 0.026. This means that, when one controls for regional differences by centering, and for other factors beyond population through sample size, the resulting estimate of the average rate of gain in productivity with increasing settlement population is within two one-thousandths of the predicted value. This result provides striking support for the model.

There is one final point that should be made about this analysis. In the contemporary world the height of a scaling relationship, captured by the intercept of the fit line, generally increases from year to year. Current theory suggests such increases are due to decreases in transport costs and increases in the energetic productivity of individual interactions. As a result, one needs to center the data by year if one wishes to use contemporary data from different years in a single analysis. In contrast, Figures 1 and 2 combine sites that date to different moments in the history of each region. There is no theoretical reason to expect that the intercept of the fit line capturing the relationship between \( N \) and \( Y \) should be static, but previous studies have not found evidence for a changing intercept over time (Ortman et al., 2015, 2016; Ortman and Coffey, 2017). The fact that the pooled analysis presented here leads to an estimate of \( \beta \) that is so close to the theoretical prediction provides additional evidence for consistency in the basic energetics of the economy in each of these regions over long periods of time. This does not mean past economies were static, but it does suggest the easiest way for societies to increase productivity is through agglomeration. This is a striking finding with obvious relevance for social policy.

This framework, and type of analysis, has been applied to a range of urban and non-urban settlement systems known through history and archaeology (Hamilton et al., 2007, 2018; Ortman et al., 2014, 2016; Cesaretti et al., 2016; Hanson and Ortman, 2017; Hanson et al., 2017; Ortman and Coffey, 2017; Altaweel and Palmisano, 2018). And it has also been applied to a range of data from contemporary urban systems (Pumain et al., 2006; Bettencourt, 2013; Lobo et al., 2013; Schläpfer et al., 2014; Bettencourt and Lobo, 2016; Mahjabin et al., 2018). So far, with allowance for a few wrinkles, the data have been consistent with specific expectations of settlement scaling models in every case. These results suggest that, at least with respect to population size, human agglomeration effects are highly predictable. This does not necessarily mean that doubling the size of a given city today would necessarily increase its per capita socio-economic rates by 16.7 percent. Indeed, there are cases from recent times where specific cities have grown substantially in population without a corresponding increase in GDP, for example (Henderson, 2003; Jedwab and Vollrath, 2015). But the theory does say that this is the average expectation. In essence SST allows one to control for agglomeration effects, thus bringing other factors that influence outcomes in specific situations into greater focus. It doesn't disregard history or context; it simply captures the physical and energetic factors that constrain the range of histories that are possible. SST only deals with the material effects of agglomeration. It does not address associated psychological or emotional effects, or indeed, any other aspects of life in cities that it would be worthwhile to know more about. But it is still a good start. Indeed, I think being able to make mathematical predictions regarding anything specific about human networks is an exciting advance.
An additional important aspect of the Social Reactors Project that archaeological evidence is not merely being used to confirm an existing theory. Rather, it is being used to expand and elaborate the theory. For example, the theory proposes that the increasing returns to scale that characterize contemporary urban systems derive from the expansion of human connectivity brought about by density. In a modern context density is a tricky concept because it is sensitive to the area over which people are counted, and when they are counted. What area should be used? What time of day? Today, the edges of built-up urban areas bear little resemblance to administrative and political boundaries, and many workers commute across such boundaries on a daily basis. As a result, it is very difficult to define the relevant spatial units, and interacting populations, that should exhibit increasing returns in contemporary urban systems. It turns out that this problem is much less severe for the smaller and simpler societies known through archaeology. In most cases, the physical settlement and its associated mixing population correspond much more closely in the archaeological record than they do today. As a result, it is actually more straightforward to test SST using archaeological evidence than it often is using contemporary data (Lobo et al., 2019).

Some may question whether the assumptions embedded in this approach—the balancing of costs and benefits, that socio-economic rates are proportional to interaction rates, the idea that interactions have energetic benefits, etc.—are appropriate. It is also reasonable to question whether the archaeological proxies used in testing these models are appropriate, and whether the data at our disposal are of sufficient quality. All of these issues aside, settlement scaling models generate testable predictions that are borne out in many datasets, using a variety of measures, from many societies, with radically different forms of political and economic organization, both past and present. So the empirical support for settlement scaling theory exists regardless of one’s prior beliefs regarding the assumptions in these models and proxies. This is important because it helps the theory stand up to cross-examination by someone who is not predisposed to accept it. To reject the theory, one needs to show that an alternative model accounts for the empirical evidence better. Urban geographers are beginning to interrogate some of the assumptions and results of settlement scaling research more closely (Arcaute et al., 2015; Depersin and Barthelemy, 2018; Keuschnigg et al., 2019). And it would be great if archaeologists contributed to this as well. This is what it will take to build an understanding of agglomeration effects that is strong, clear, and specific enough to guide us into the future.

A NEW KIND OF RELEVANCE

In presenting the example of SST, I do not mean to suggest that the only way archaeology can achieve practical relevance is through the development of explicit formal models. Indeed, in many cases medical researchers show that specific medicines have quantifiable therapeutic effects even when they can’t explain the mechanisms behind them. And there are examples of this kind of logic being applied to archaeology. As an example, Ingram (2015) recently investigated human vulnerability to drought by comparing paleoclimate records with measures of settlement instability for large numbers of settlements located in a variety of ecological settings. Among other things, his analyses found a strong relationship between drought and migration that was insensitive to the proximity of residents to a perennial water source. Additional studies of specific situations like this clearly have the potential to guide future decisions, even in the absence of a formal model.

Regardless of how well SST stands the test of time, I hope this example successfully illustrates that it is possible to build predictive knowledge of human affairs that incorporates but also transcends the archaeological record. The process has just barely begun, but if we believe the archaeological record is at least partly systematic, that human behavior is at least partly predictable, and that scientific reasoning can be employed to improve the human condition overall, this seems like a very good thing to incorporate into an expanding scope of archaeological practice.

Such research is challenging. It requires careful observation of the phenomenon to be explained, definition of key concepts and relations, formulation of theory and models, painstaking work to compile the relevant evidence for testing, careful analysis of the data, and critical evaluation of the entire process. But it is not impossible. The basic logic and analytical procedures for testing such models are already part of the standard training of archaeologists. Middle-range theory continues to provide a basis for constructing valid proxies for human behavior that researchers outside of archaeology will find relevant. Dramatic recent expansion in the ability to collect data on contemporary human behavior is stimulating exciting developments in other social sciences. And the example of settlement scaling theory shows that it is possible to develop predictive theory that is amenable to empirical testing and applies as well to societies known through archaeology as it does to societies that can be observed directly today. As a result, knowledge of the social reactor process emanating from archaeological research should be relevant for urban science and urban policy. I see no reason why this could not also be done for a wider range of contemporary social issues with additional effort, and with more interdisciplinary collaboration.

There is great social benefit in all the things archaeologists do—from heritage management to museum exhibits, cultural tourism, advocacy, historical reconstruction, traditional synthesis, and cross-cultural analysis. We should be proud of everything we do, and keep on doing it. The purpose of this paper has been to suggest that in addition to all this we can and should strive to expand the contemporary relevance of archaeology such that the results of archaeological research can help us make informed decisions in charting a better future as we confront today’s challenges. The archaeological record is the richest and most extensive source of information on human social experience we have. In coming years, I hope more of us will work to develop this record to its full potential.
DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/Supplementary Files.

AUTHOR CONTRIBUTIONS

SO conceived of and wrote this paper.

FUNDING

Publication of this research has been supported by a grant from the James S. McDonnell Foundation (#220020438).

REFERENCES


ACKNOWLEDGMENTS

I wish to thank Tim Kohler, Mike Smith, Keith Kintigh, Jeff Altschul, and the article reviewers for many helpful comments on previous versions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdigh.2019.00016/full#supplementary-material

Data Sheet 1 | House counts and total house areas for archaeological settlements in five New World societies.


Wright, H. T., and Johnson, G. A. (1975). Population, exchange and reproduction is permitted which does not comply with these terms.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Quantitative Methods for the Comparative Analysis of Cities in History

Luís M. A. Bettencourt 1,2,3* and José Lobo 4,5*

1 Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, United States, 2 Department of Ecology & Evolution, Department of Sociology, University of Chicago, Chicago, IL, United States, 3 Santa Fe Institute, Santa Fe, NM, United States, 4 School of Sustainability, Arizona State University, Tempe, AZ, United States, 5 ASU–Santa Fe Institute Center for Biosocial Complex Systems, Arizona State University, Tempe, AZ, United States

Comparative studies of cities throughout history are one of the greatest sources of insight into the nature of change in human societies. This paper discusses strategies to anchor these comparisons on well-defined, quantitative and empirical characteristics of cities, derived from theory and observable in the archeological and historical records. We show how quantitative comparisons based on a few simple variables across settlements allow us to analyze how different places and peoples dealt with general problems of any society. These include demographic change, the organization of built spaces, the intensity and size of socioeconomic networks and the processes underlying technological change and economic growth. Because the historical record contains a much more varied and more independent set of experiences than contemporary urbanization, it has a unique power for illuminating present puzzles of human development and testing emergent urban theory.

Keywords: urbanization, scaling, zip's law, economic growth, data

INTRODUCTION

Cities have always held a special fascination to any scholar of human societies. Coincident with the advent of the first cities, we observe the appearance of many technologies and adaptations that, in different forms, are still with us today (Adams, 2005). Thus, the experience of living in cities (Wirth, 1938; Lees, 2015) provides a general conducting line throughout history, connecting common phenomena across different societies and thus also identifying features that are truly contextual.

Performing comparative analyses of different societies is always an exercise fraught with challenges. There is the empirical challenge of identifying cultural, social, political and economic traits, which can be measured in very different settings. But there is another difficulty when doing comparative analysis which habitually goes unnoticed. The identification of common traits is often conditioned on performance measures, such as rates of economic growth or energy use per capita, which convey a sense of what today we find important (Mcfarlane, 2010). Assessing the nature, and even the quality, of ancient societies can easily be biased by using the socioeconomic experience of today's high-income nations and their recent history. Is evidence for improvements in diet or material conditions in ancient societies to be disregarded because these same societies did not experience high (by today's standards) output growth rates? Such a stance crudely disregards many of the extraordinary adaptations and inventions—social, cultural, and technological—of earlier societies. Alternatively, conditioning on environmental stewardship and sustainability leads to the...
opposite conclusion, ranking smaller scale societies that had less impact on their immediate natural environments as having higher quality than most recent societies.

There is however an alternative to such approaches, which starts with much more basic but also more pervasive features of any settled human society (Bettencourt, 2013; Ortman et al., 2014). A number of recent new ideas, supported by extensive empirical analyses, point to certain quantitative comparisons of basic general quantities that may shed light on a number of key puzzles about the organization, sociality and capacity for adaptation of past urban societies (Bettencourt et al., 2007; Fletcher, 2011; Bettencourt, 2013; Ortman et al., 2014, 2015, 2016; Cesaretti et al., 2016; Hanson and Ortman, 2017; Ortman and Coffey, 2017). Such puzzles include the relative size, structure and flows between settlements in the same polity, the nature of socioeconomic networks in cities, the spatial organization of settlements, and the nature of change and adaptation in these systems, including processes of economic growth (Economic growth is here understood to be simply an increase, from one period to the next, in a society’s material output). What is most important to capture through such comparisons, in our view, is how different societies deal with general problems affecting them all, including energy and resource extraction, and the organization of their socioeconomic networks over space and time (Bettencourt, 2013; Morris, 2013; Ortman et al., 2014).

As we look back at history from a modern perspective, shaped by an urban planet with large human population and fast economic growth and technological change, these puzzles become especially poignant: Are pre-industrial societies fundamentally different in the way people lived and interacted? Or are these differences primarily connected to issues of scope, scale and technology? Can we identify, clearly and empirically, lines of continuity and divergence in the structure and dynamics of urbanizing societies?

These puzzles cannot be answered simply by using the present as the baseline for comparison: what is needed is a framework that makes comparison between the experiences of the past and life in the present conceptually coherent and empirically consistent (Bettencourt, 2013; Ortman et al., 2014). Here, we explore three strategies for quantitative analysis of settlements throughout history. We discuss how these are undertaken methodologically and their promise for generating a more integrated understanding of our social history as well as an appreciation of each society in its own context. A comparison of the past and the present that is based on fundamental processes and features makes it intelligible to use the past and present to discern what the future might be like.

RESULTS

Because we are asking for quantitative ways to perform comparative analysis of cities in history we need to obtain data that are consistent across places and times. This remains a challenge, not only because empirical evidence in the archeological record is sparse and mostly associated with durable materials, but also because methods and definitions have naturally varied between many different communities, each dedicated to different periods, using different methods of analysis, etc (Kintigh et al., 2014).

Thus, to go forward and attempt any reasonable synthesis, simplicity and clarity are paramount. Simple quantities such as the area of a settlement, its putative population count (based on independent measures, such as room counts or amounts of debris), and perhaps other basic quantities related to public spaces or monument construction are usually available through the material record, and have now been measured in several instances (Bettencourt, 2013; Ortman et al., 2014, 2015, 2016; Ortman and Coffey, 2017). The analytical advantage of these quantities is that they are reasonably objective and salient features of any human settlement, while leaving plenty of room for varying cultural, political, and economic features of different societies (McFarlane, 2010; Lees, 2015).

For simplicity then, we ask below what we can be inferred from fairly sparse data records, where only a few variables (one, two,...) are available for each site. This approach also allows us to connect to well-known traditions in history, demography and geography (Fujita, 1990; Bairoch, 1991; Zipf, 2012; Morris, 2013; Ober, 2016), before we attempt to take longer steps toward the end of the paper.

One Variable: Demography and the City-Size Distribution

Perhaps the most established way to characterize an urban system quantitatively is by analyzing the statistics of settlement sizes, or equivalently testing the “rank-size” rule (Henderson, 1974; Fujita, 1990; Bairoch, 1991; Zipf, 2012). This is the simplest of all tests of any quantitative expectation for cities. It requires data on only a single variable, such as the population of each settlement. For this reason, studies constructing the settlement size distribution for many societies are numerous and have been undertaken for decades (Bairoch, 1991; Gabai, 1999; Zipf, 2012; Swerts and Pumain, 2013). In many archeological applications, population is replaced by more directly observable proxies, such as the settlement’s area.

The simplest expectation for the rank-size rule (also known as Zipf’s law Krugman, 1996; Zipf, 2012 states that, when cities are rank-ordered from largest (rank = 1) to smallest (rank = number of cities in the system), the size of each city is simply inversely proportional to its rank:

$$\text{size}(\text{rank}) = \frac{\text{size}_{\text{max}}}{\text{rank}^z},$$

where \(\text{size}_{\text{max}}\) is the size of the largest city and \(z\) is an exponent. The standard rank-size rule applies for \(z = 1\). This is equivalent to the probability distribution of city sizes taking the form

$$P(\text{size}) = \frac{P_0}{\text{size}^{1+z}},$$

where \(P_0\) is a normalizing constant, so that the probability integrates to unity.

Much has been made of the shape of the city size distribution and its meaning. The common exercise deals with the estimation...
Mechanically, the size of each city measures simply the integrated growth (including periods of decline) over its history, which is essentially a measure of its demographic average growth rate over a long period of time. The simplest rank-size rule states, from this perspective, that all settlements grow at the same rate (Gabaix, 1999) (if they were created at the same time), another approximate statistical regularity known as Gibrat’s law. Note that this does not have to mean that demographic growth rates are the same for all cities at all time, but simply that over a long period of time these rates converge to the same number, presumably as the result of balances between births, deaths and migration between these towns and cities.

Thus, by comparing the relative size of different settlements, historians and archeologists should be asking whether these were part of the same “demographic” system, connected by mutual migration flows and other networks of exchange and trade. If so, observing something close to the settlement size distribution predicted by the rank-size rule would imply the same average population growth rates for all places, big and small. Then, for example, if mortality rates were higher in larger cities, this would imply a correspondingly larger rate of immigration from smaller places to larger cities (Dyson, 2011; Bocquier and Costa, 2015).

If it is possible to measure the size distribution of the same settlements at two or more times, then we can moreover compare their relative growth rates during the intervening periods, giving us an empirical basis to rank their relative (demographic) success.

Some additional issues are worth flagging here. For human settlements, spatial areas are not typically proportional to population sizes. How physical space is used socially can be modulated by cultural and physical infrastructure, as well as by technology (Wirth, 1938; Adams, 2005; Bettencourt, 2013). Furthermore, areas can also be measured in different ways, as the surface within the convex hull of the settlement's putative boundaries $A$, or as the actual built up area of buildings, streets and other structures, $A_n$. We discuss these issues next.

Two Variables: Settlement Scaling, Density, and Agglomeration

What was life like in the ancient city of Ur? Or in the great city of Teotihuacan? We will never really know 150 for sure, of course. One of the main objectives of archeological and historical research is to reconstruct what social and economic life might have been like from fragmentary information, much of it about the 152 built environment. This is a very difficult type of inference that requires a testable theory of how properties of social and economic life relate to variations in specific characteristics of the built environment.

Settlement scaling theory attempts to do precisely this (Ortman et al., 2014, 2015, 2016; Ortman and Coffey, 2017). Developed originally to explain urban scaling properties in contemporary cities (Bettencourt et al., 2007; Bettencourt, 2013; Ortman et al., 2014), its ingredients are very general leading to the exciting prospect of the application of its core ideas to settlements in history. The empirical observations on which it is based, as well as its core models, indicate that several basic social economic and infrastructural properties of settlements are interrelated, and
can thus be predicted on the basis of comparative analyses of their built environment and estimates of their population size. Empirically, scaling analysis is also very simple, requiring only pairs of variables for each settlement, and the analysis of a familiar xy plot (Figure 2).

If we have two variables for each settlement we can ask for example, how does their built-up area depend on their population size: all we have to do is plot one quantity against the other. The answer tends to be non-linear, but well-described by scale invariant functions (power-laws), such as

$$A(N, t) = a(t)N(t)^\alpha.$$  \hspace{1cm} (3)

This can be made linear by a simple transformation to logarithmic variables, or a loglog xy plot (Figure 2).

Theoretical considerations derive the values for the prefactor \(a(t)\) and the exponent for area as \(2/3\ \alpha\ 5/6\), depending on the type of settlement and how area is measured (Bettencourt, 2013; Ortman et al., 2014). These expectations are confirmed by empirical analysis of many settlement systems, including in the pre-Columbian Basin of Mexico (Ortman et al., 2015, 2016), classical Rome (Hanson and Ortman, 2017), Medieval Europe (Cesaretti et al., 2016), and of course contemporary urban areas (Bettencourt et al., 2007; Bettencourt, 2013).

The same theoretical framework predicts scaling relations and exponent numerical values for many other quantities (see e.g., Figure 2), including the number of socioeconomic interactions in a settlement, its division of labor, its rate of socioeconomic production and many detailed characteristics of the built environment, such as street length and width and associated transportation costs (Bettencourt, 2013).

In this way, a very straightforward two-variable scaling analysis can reveal commonalities of settlements as socioeconomic networks self-consistently embedded in built spaces. An expansion of this type of analysis to other settlement systems promises to reveal common quantitative patterns of basic settlement organization and socioeconomic capacity in societies through space and time.

It is also from the perspective provided by these observations and associated theoretical frameworks that we may appreciate any exceptions. For example, an interesting set of questions has been raised by Fletcher about “low density urbanism” (Fletcher, 2011), specifically in the context of Mayan settlements and Angkor Wat, which appear to show an expansion of their area with population with an exponent, \(\alpha > 1\). Therefore, such settlements would become less dense the larger they are, not realizing agglomeration effects typical of other cities. A similar, but perhaps more expected pattern also applies to mobile hunter-gatherer camps, but with greater variability. These patterns also vary in time in specific ways, to which we now turn.

Technological and Economic Change

In modern societies, cities have been a necessary condition for economic growth (Jones and Romer, 2009).

We say necessary because the existence and expansion of cities is not always sufficient for income growth at the national level: there are many episodes, some shorter and some longer, of urbanizing societies experiencing no (economic) growth (Inoue et al., 2015). Nevertheless, the association between higher levels of urbanization and larger GDP per capita is one of the strongest empirical results in studies of economic change and international development.

Much work has been done to try to elucidate this connection and better understand the mechanisms of technological change and economic growth generated by urban environments (Lucas, 1988; Jones and Romer, 2009). However, if the judgment of success is predicated on creating quantitatively precise growth rates, it remains fair to say that the problem is not yet well-understood.
Many studies in economic history have also shed light on the circumstances that led to sustained economic growth after the industrial revolution, calling our attention to macroeconomic factors such as the availability of energy on a large scale, political and economic institutions, and the advent of modern science (Morris, 2013). The study of socioeconomic development in the past has also highlighted the role of urbanization (Algaze, 2008; Cowgill, 2015; Ober, 2016; Harper, 2017; Manning, 2018) as have historical experiences of urbanization without growth (Jedwab and Vollrath, 2015). As useful as detailed case studies and historical examinations are, comparative analyses have been hampered partly because of a perceived lack of common empirical evidence within regions and across eras and geographies.

There are, however, a set of facts that may be useful for framing the study of urbanization’s role across time: i) sustained economic growth is a system’s level property (see Figure 2 for China); ii) growth volatility reduces rates of economic growth; iii) very small rates of systemic economic growth are not perceptible over a human lifetime; and iv) the accumulation of material wealth resulting from low-levels of growth are vulnerable to exogenous shocks (such as disease or changes in climate). As a consequence, growth can go unnoticed and remain accidental. This is not to say that people were not keenly aware of times of prosperity or famine, resulting from conquest or good harvests, it simply means that the concept of long-run intensive economic growth would have been very hard to perceive and nurture in pre-industrial societies.

The first point may not be obvious as we often think of rich and poor settlements, even within the same nation or polity. It is, however, generally true that the type of sustained and fast economic growth observed in modern settlement systems
is a system level property (so that information, ideas, resources and individuals can flow among settlements), with all cities experiencing about the same annual rate of growth over long periods of time (see Figure 2).

The happy consequence of this observation is that studying systemic economic growth in history may require only a number of point assessments, which should agree in magnitude whether they were measured in small towns or larger cities. This also means that golden ages often associated with large cities, such as classical Athens or Rome, whether triggered by a technological innovation or by conquest and theft, may not be sustainable unless they induce economic growth across their settlement systems (Ober, 2016). This means, for example, that we should see the living experience of primary producers living in small settlements change so as to enjoy some of the products of large cities and vice-versa in a virtuous cycle of exchange and common development. We know of course that prior to the industrial revolution such periods, if they existed at all, were not associated with large growth rates, and were typically localized in space and time.

The second and third properties of economic growth follow from its character as a stochastic (fluctuating) process. This is a very general feature of collective dynamics of growth, from population biology to financial markets (Bettencourt, 2018). Without going into detailed models for those contexts, quantities such as the resources available to a society (‘wealth’) are expected to grow approximately as

\[
\frac{dr(t)}{dt} = (\eta + \epsilon)r(t),
\]

where \(\eta\), \(\epsilon\) are the average growth rate (an approximate constant in time, say 1% a year) and the corresponding stochastic variations, respectively.

Writing the variance of \(\epsilon\) as \(\sigma^2\) (also known as the square volatility) allows us to integrate the equation in time to give

\[
\ln \frac{r(t)}{r(0)} = \left(\eta - \frac{\sigma^2}{2}\right)t + \sigma\sqrt{t}\xi(t),
\]

where \(\xi(t)\) is an approximately normal variable with zero mean and unit variance. The actual growth rate \(\eta - \frac{\sigma^2}{2}\) that results is the geometric mean (not the arithmetic mean!) of growth rates, as is well-known in population biology. This is reduced from the average growth rate by a term proportional to the square volatility, \(\sigma^2\), that is half the variance of the growth rate, due to its fluctuations over time. Thus, high variability can render any small growth rate zero or even negative (see Figure 3). This means that innovations to reduce instability in the economy are, in the beginning, almost as important as having a positive growth rate in the first place.

With all that said, the final argument we wish to emphasize here is that the growth rate for any preindustrial economy over any extended time period (say decades) was likely very small. Figure 3, based on lead emissions, suggests a value of about 0.17%, certainly lower than 0.3% a year. This translates at the most into a doubling time for the economy of 240 years. This time scale is too long to be felt by anyone—on average at least—in their own lifetimes. Thus, even if slow economic growth was present in preindustrial societies, it was likely too slow for its society to become conscious of it and take measures that could sustain it. The perception would then be one of effectively zero growth, where any positive period would be quickly reversed by fluctuations.

Even if the change in material output of societies in the past had been exponential in nature, the accumulation of wealth could have been greatly set back by disease, climate change or war. And
even if the underlying social processes by which agglomerated populations learn, innovate, and become more productive are the same across eras, societies abilities to deal with the plague or sharp reduction in rainfall are importantly determined by science and technology. In conclusion, processes of human development and economic growth recognizable to us today were probably at play throughout history, and certainly in most urban societies. However, even in the best of times such rates of change may have been too local, too volatile and too short-lived to be acted upon and sustained, intentionally, over the long term. The search for some of the tell-tale signs of these episodes, especially in the systemic change in living conditions across settlement sizes may give us precious new insights into the actual time dependent variability of these effects, and on the human experience in cities during long periods of very slow growth and even decay.

DISCUSSION

The history of cities presents us with a bewildering variety of social, economic, political and cultural ways in which human settlements can exist. Making sense of this variation, while at the same time extracting what may be essential across time and space, is a necessary goal not only for a “science of cities” but for a science of human sociality. We have shown how an approach to comparative analysis based on common but determinant variables for human settlements—including population, area and measures of the built environment- has the power to support an analytical narrative relating the earliest settlements in history to contemporary cities and presumably their future forms.

The ambition to develop a theoretical and empirical basis for the study of human settlements through history may invoke in the reader common criticisms of any cross-cultural comparative analysis, specifically that some societies will be judged to be better or worse, and that contemporary high-income societies along with their economic and political systems will be used as standards for evaluation. None of this follows from the strategy proposed here, except the fairly mechanical features that some societies are larger than others—in both their creative and destructive capacity—and so must possess knowledge of, and access to, different types of resources, in ways that are sometimes sustainable and sometimes exploitative. Only by learning formally about this variation can we come to appreciate the range of the human experience in cities.

The approach proposed here then simply connects social and cultural life to some of its most basic material underpinnings, common to all societies in all places. This includes the fact that people exist in space and that their interactions must be structured over space and time in ways that must be compatible with their collective socioeconomic capacity. Evidence from historical and archeological sources have the singular potential to illuminate these issues in ways that contemporary evidence cannot.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.dropbox.com/sh/xuuuy90edvlekbp/AADhWgmmVuiMOU2txx-Kc2aNa?dl=0.

AUTHOR CONTRIBUTIONS

LB has conceptualized the study and analyzed the data. LB and JL wrote the manuscript.

FUNDING

This work is partially supported by OPEN-AIRE funding linked to the Marie Curie Project Past-people-nets 628818, conducted by Francesca Fulminante (2014–2016). Publication of this research has been supported by a grant from the James S. McDonnell Foundation (#220020438).

ACKNOWLEDGMENTS

We are thankful to the speakers and participants of two workshops, Urbanization without Growth in the Ancient World? (at the British Academy in Rome, July 2017) and at the Roman Archeology Conference (RAC/TRAC) 2018, where some of the material presented here was originally developed.

REFERENCES


Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bettencourt and Lobo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Ancient City, Universal Growth?
Exploring Urban Expansion and Economic Development on Rome’s Eastern Periphery

Matthew J. Mandich*

International Society for Archaeology, Art and Architecture of Rome (ISAR), Rome, Italy

This article investigates the urban expansion and economic development of ancient Rome through the application of models and theories originally designed for the study of contemporary cities. While the growth of ancient settlements is often difficult to track and analyze, archaeologically observable changes in land use can be read and interpreted as a function of broader economic oscillations over the *longue durée*. By re-examining the available archaeological and textual evidence pertaining to land use change on Rome’s eastern periphery this article demonstrates how the frameworks selected can be successfully appropriated via a narration of Rome’s urban transformations from the mid-Republic to the later Imperial period. The ultimate goal is to determine if the patterns of urban expansion identified in modern cities also existed in ancient Rome. The findings provided have the potential to produce rich insights on the dynamics of urban and economic growth across time and geographies, thereby opening the door for new and further studies.

Keywords: Rome, Roman archaeology, Roman topography, economic geography, fringe belts, location theory, settlement scaling theory

INTRODUCTION

“If anyone wishes to estimate the size of Rome by looking at these suburbs he will necessarily be misled for want of a definite clue by which to determine up to what point it is still the city and where it ceases to be the city; so closely is the city connected with the country, giving the beholder the impression of a city stretching out indefinitely.”

-Dionysius of Halicarnassus, 1 Roman Antiquities 4.13.4

Defining Rome’s urban area was no easy task in antiquity and it remains difficult for today’s archaeologists and topographers to track and determine its ancient urban extent(s). Although the mid-Republican city was demarcated by an 11 km circuit wall, constructed from circa 378–353 BC, 2 these fortifications should not be seen to represent Rome’s true urban extent at that time since they encompassed a space (c. 427 ha) much larger than the inhabited area, likely containing swaths...
devoted to protected agriculture (Livy\textsuperscript{3} 2.11.3, 6.32.1; Coarelli, 1988, p. 323–330). However, as the City’s population grew considerably from the Fourth to First centuries BC land beyond the walls was put into use at an ever-increasing rate (Lo Cascio, 2010, p. 30–38). Certainly, by the reign of Augustus—the period in which Dionysius of Halicarnassus was active (c. 20 BC)—a denser network of buildings stretched well past the city walls forming what could today be considered somewhat of a conurbation.

This continuously built area, or continentia aedificia, as it was referred to in legal texts, already existed beyond the walls in the time of Caesar and its continued expansion presented multiple cadastral challenges (e.g., Cicero\textsuperscript{4}, Ad Atticus 13.33a; Suetonius, Caesar, 44; Lex Iulia Municipalis\textsuperscript{5}; Digesta\textsuperscript{6} 50.16.87, 50.16.139, 50.16.147). Augustus’ reorganization of Rome into 14 regions (or regiones) was designed to address many of the administrative and safety issues inherent to the expanding cityscape, including the management of extramural space (Suetonius, Augustus 30; Frezouls, 1987, p. 375). As can be seen in Figure 1, six of the 14 Augustan regiones encompassed land outside the old (Servian) city walls, seemingly in areas where building was most dense and/or conspicuous—i.e. in the Campus Martius and along

\textsuperscript{4}Cicero (Translated by H. G. Hodge 1927) Ad Atticus. London: Harvard University press.
\textsuperscript{5}Lex Iulia Municipalis (Translated by A. C. Johnson, P. R. Coleman-Norton, F. C. Bourne, C. Pharr 1961). Austin: University of Texas Press.
certain thoroughfares to the south and east. It is often overlooked that these extramural regions had no external boundaries prior to the construction of the Aurelian Wall, begun in AD 271, and they may have been designed to grow without constraint. Indeed, as Quilici (1974) and Frezouls (1987) have observed, Augustan Rome could certainly be considered an “open city,” spreading out like “tentacles” along its numerous paved viae. Although traditional and administrative boundaries, such as the pomerium (Rome’s oldest religious boundary) and the octroi (customs) boundary were also used to delimit urban and extra-urban spaces and activities, in some cases affecting the topography of the City, these borders were not designed to adapt to or keep pace with Rome’s rapid expansion and should not be understood as accurate representations of its true urban extent (see Palmer, 1980; Goarelli, 1997; Andreussi, 1999; Giardina, 2000).

Instead, the rough edges of Rome’s expanding, extramural built area remain the most realistic representation of its size from the late Republic through the Imperial period. Yet, because the ephemeral limits of the continentia aedificia have been difficult to locate and track, the significance of its advancement has been largely ignored. However, by taking an interdisciplinary approach, employing models and theories from urban morphology, economic geography, and complexity science, ancient Rome’s urban expansion can be followed and linked to cycles of economic growth and decline. In particular, bid-rent theory, the fringe belt model, and settlement scaling theory are useful for tracking Rome’s physical growth and interpreting it as a function of concomitant economic development. To demonstrate how these frameworks can be successfully applied, archaeological and textual evidence pertaining to Rome’s eastern periphery is used to analyze land use patterns beyond the circuit wall from the mid-Republic (Fourth century BC) to the late Imperial period (Third century AD). How the patterns observed compare to those seen in modern settlements is key for determining if Rome was expanding (at least for a period) like a post-industrial city.

MATERIALS AND METHODS
Understanding the Implications of Urban Growth: Bid-Rent Theory and Settlement Scaling Theory
The concept of rent, as defined by Ricardo (1817, p. 34–35), is the “compensation that is paid to the owner of land for the use of its original and indestructible powers”. According to Ricardo (1817, p. 34–35), rent exists “because land is not unlimited in quantity and uniform in quality”; that is, as soon as land of a more marginal quality is put into use (often due to population pressure), rent will immediately commence on land of higher quality. This concept of economic rent was central to Von Thünen (1826) seminal “Isolated State” model, which focused on the spatial distribution of agricultural practices and land use around an isolated city (i.e., market center). Although purely hypothetical, it highlighted distance-based agricultural activities, taking into account production costs, transport costs, and profit maximization to determine a more nuanced version of land rent. The resulting situation is a rent gradient in which rents decrease with distance from the market (settlement/city) as transportation costs increase, creating a series of concentric zones in which particular types of agriculture are practiced (Figure 2).

Yet, how and why land uses locate in an urban setting is considerably more complex than in a rural one since space is more restricted and land is assigned to the highest bidder—i.e., the individual or institution willing to pay the most rent. Burgess (1925) was one of the first to examine land rents in an urban setting and his “concentric circle model,” which was designed to explain and predict the distribution of social groups within a city (i.e., Chicago), showed how patterns of residential land use emerged due to multiple competing factors (Figure 3). To better predict this variability, Alonso (1964) devised “bid-rent theory” to examine the location of multiple types of land use in an urban setting (e.g., commercial, residential, institutional). His theory used a detailed mathematical framework to produce “bid-rent” (or bid price) curves that vary based on the type of land use analyzed (i.e., the type of bidder), in addition to accounting for non-economic factors, such as “trade-offs.” While a simplified version of the bid-rent model indicates differing rent gradients will form concentric zones around a center, each featuring a

![Figure 2](https://example.com/image2.png)

**FIGURE 2** | The Von Thünen model with multiple variables introduced. Segment I depicts the model in its undistorted (hypothetical) shape; Segment II shows how differences in the quality of soil can transpose the cultivation of certain products outside their “expected” zone; Segment III illustrates the influence of a second market; Segment IV shows the effect of peasant farmers on the model as they tend to grow products based on their personal needs rather than market principles; Segment V introduces a combination of variables including roads, rivers, minor centers, and uneven geography, which is closest to the real-world scenario (redrawn after Pred, 1967, p. 26, Figure 8.6).
dominant land use given the desirability of the location to the bidders (much like the Von Thünen and Burgess models) (Figure 4), Alonso’s original model is much more detailed since it shows how land rents are affected by complex factors.

For example, when analyzing residential bid-rent curves, Alonso’s model illustrates how the steepness of rent gradients (and the location of property types) is affected by population growth, transport technology, and even the purchasing power of the individual (Figure 5). As the figure shows, population growth forces residential rent gradients up due to higher demand, but higher incomes lessen the steepness of residential bid-rent curves. This means that wealthier individuals tend to live on the periphery of cities because they can spread the cost of commuting over larger sites (i.e., a “trade-off”—in this case, exchanging cheaper transport costs for more space) (Alonso, 1964, p. 106–109). Improvements to transport technology should also reduce the steepness of bid rent curves more generically since they lower overall transportation costs; however, as Alonso (1964, p. 112–113) observed, such improvements make residential land prices on the periphery of cities higher (due to competition between bidders) since a reduction in transport costs allows the same level of satisfaction to be achieved at a further distance.
from the center on a plot with more space. As such, the vicinity of a road works to raise land values considerably, as rent gradients are both elevated near main roads and extend further along them, thereby extending the limits of the settlement (Alonso, 1964, p. 141–142) (Figure 6). In terms of economic growth, Alonso (1964, p. 114–115) concluded that periods featuring simultaneous population growth, transport innovations, and rising per capita incomes are indicative of rapid economic development, and that such a combination should result in a slower rate of outward urban expansion, leading to simultaneous densification.

This conclusion has recently been wholly confirmed through quantitative analysis provided by the emerging framework of settlement scaling theory (SST), which plots a range of factors against urban population numbers to determine how attributes of settlements change (or scale) with settlement size. Using data obtained from modern cities, the resulting calculations have shown that as the population of a city doubles, the necessary infrastructure only increases by about 85% (i.e., $\beta = 0.85 < 1$), while proxies for socio-economic activity increase by about 115% ($\beta = 1.15 > 1$) (Bettencourt et al., 2007). Thus, as cities grow exponentially (double in population) their infrastructure scales sub-linearly with population as a spatial economy of scale (they become denser), while their socio-economic outputs and per capita growth scale super-linearly, showing increasing returns to scale (they become more productive) (Bettencourt et al., 2007, p. 7303; Bettencourt and West, 2010, p. 912–913) (Figure 7).

The existence of these scaling relationships reveals two distinct, often competing, aspects of urban growth, with one based on materials, infrastructure, and efficiency (sub-linear—economies of scale), and the other on social interactions, innovation, and wealth creation (super-linear—increasing returns to scale) (Bettencourt et al., 2007, p. 7303). While sub-linear scaling relationships are often associated with “extensive” economic growth driven by aggregate growth (i.e., more input equals more output), super-linear relationships are associated with “intensive” economic growth defined by technological innovation and/or divisions of labor that produce rising per capita income (i.e., inputs used more productively to create greater outputs) (Lal, 1998, p. 19–26).

Similar scaling relationships have been found to exist in settlement data from pre-modern contexts, including the pre-Hispanic Basin of Mexico, the Inca Empire, Medieval Europe, and the Roman Empire, indicating that these correlations should be attributed to the processes of human agglomeration rather than specific institutions or technologies (e.g., Ortman et al., 2015, 2016; Cesaretti et al., 2016; Hanson and Ortman, 2017). So far, only sub-linear scaling relationships have been found to exist in settlement data from the Roman world; however, this does not preclude the existence of super-linear relationships in a Roman context since increasing socio-economic returns should go hand in hand with demographic growth, technological innovation, and sub-linear areal expansion (i.e., increasing density and infrastructural efficiency) (see Mandich, 2016, p. 194–196).
Yet because ancient settlements were sensitive to numerous exogenous and endogenous factors, often featuring complicated non-linear trajectories that differed by settlement, region, and period, a finer-grained approach that accounts for both urban growth and contraction is necessary to properly track system change over time.

Tracking Urban Expansion: The Fringe Belt Model

The fringe belt model is a morphogenetic approach used to analyze the physical and economic evolution of settlements through an examination of land use change on the urban periphery. Because this framework focuses on observable patterns of change, rather than the decision-making processes behind them, it is particularly well-suited for studying ancient settlements. How and why fringe belts form, how they can be identified (especially in the archaeological record), and how they are modified or internalized is key for understanding the processes of urban expansion and tracking the advancement of Rome’s *continentia aedificia*. Furthermore, as fringe belt formation and modification processes are linked to economic cycles (“booms” and “troughs”), this framework is especially useful for pinpointing periods of economic growth and change.

The study of fringe belts first began in Germany, where the field of urban morphology may also be said to have originated in the late Nineteenth century. Louis (1936) was the first to recognize the existence of urban fringe belts (originally called *Stadtrandzonen*) in his historico-geographic study of Berlin. In this work he identified a number of land-use zones, or urban fringes, that had developed beyond urban boundaries and were later encompassed by subsequent building activities. Conzen (1969, p. 125) extended the work of Louis in England, concluding that the outward growth of the urban fringe was dictated by periods of acceleration, deceleration, and standstill associated with building booms and troughs linked to fluctuations in population, economic development, and innovation (similar to the annual growth of a tree trunk). Thus, for fringe belts to emerge, clear pulsations in urban growth tied to economic cycles are needed (Whitehand, 1987, p. 76–83; Conzen, 2009, p. 33).

Fringe belts first tend to form around clear “fixation lines,” which are often natural obstacles (e.g., rivers, hills) or defunct city walls (Conzen, 1969, p. 125; Carter and Wheatley, 1979). While it is not always easy to identify fringe belts in morphological studies of urban areas, it is achievable, since they are composed of several distinct land uses or site types, often of an extensive nature. These include, but are not limited to [list derived from Thomas (1974), Barke (1976), Conzen (2009)]:

- **Open space**: cemeteries, public parks, market gardens, allotments
- **Institutions**: religious retreats, military barracks, community buildings, hospitals, waste disposal/dumps
- **Industry**: warehouses, quarries, manufacturing, slaughter houses
- **Residential (low density)**: villa/country estates, sub-standard dwellings
- **Recreation**: sports grounds, riding grounds, hunting/fishing areas, resorts, taverns.

As fringe belts tend to form gradually, either around a fixation line or on the edge of the halted urban periphery, the fringe belt *formation stage* can be broken down into several sub-phases (Conzen, 2009, p. 33–34):

- **Fixation phase** (incipient character associated with a fixation line)
- **Expansion phase** (pronounced character)
- **Consolidation phase** (dominant character).

The fringe belt *formation stage* is then followed by the *modification stage*, which can produce several different outcomes depending on the intensity of urban expansion occurring...
(Figure 8). First, the belt may be encroached upon heavily by redevelopment (predominantly for intensive residential or commercial uses) resulting in fringe belt reduction, where the fringe belt breaks apart, becoming smaller, and less coherent. Fringe belt reduction tends to occur when the core of the city (or central business district—CBD) expands, enveloping the previously formed fringe belt. This type of urban expansion is associated with the process of land use succession (a concept borrowed from Ecology), which is the tendency of an inner zone to expand in size via the “invasion” of an adjacent outer zone (Burgess, 1925, p. 50–52; Barke, 1976). Second, due to intense redevelopment, certain fringe belt activities or sites may be forced or coerced to migrate further afield in what is known as fringe belt translation. Third, the original use of the fringe belt may survive unchanged, forming a “relic” fringe belt that attracts the same or similar land uses as the built area continues to expand beyond it (see Whitehand, 1987, p. 83–93; Barke, 1990, p. 283; Conzen, 2009, p. 33–34).

Recent research and empirical evidence has shown that fringe belts are a widespread morphological phenomenon, occurring across every populated continent, and at various geographical scales (Conzen, 2009, p. 35–37). This has given rise to a number of questions concerning how cultural contrasts, political power structures, and the size and age of cities affect fringe belt formation and modification, as well as how these features can be studied, modeled, and measured within a more uniform methodology (Conzen, 2009, p. 46–50). Although the fringe belt model has long been known in the fields of urban morphology and geography, archaeologists have yet to test or employ this framework on an ancient city. Thus, in the following section the fringe belt model is used to examine archaeologically detectable land use patterns from Rome’s eastern periphery to determine if fringe belts existed in ancient Rome and if the same processes of urban expansion occurring in modern cities were also unfolding in the ancient world.

RESULTS: EXAMINING ANCIENT ROME’S EASTERN PERIPHERY

Rome’s Esquiline and Caelian hills were considered peripheral zones from an early period as they originally served as funerary areas for the budding proto-urban settlement (Ninth to Sixth centuries BC) (Albertoni, 1983; Colonna, 1996). These hills were incorporated into the settlement proper when it was divided into four regions, either under Servius Tullius or in the early Republic (Sixth to Fifth centuries BC) (Dionysus of Halicarnassus, Roman Antiquities 4.13–14; Pais, 1905, p. 140). Following this quadripartite division a massive rampart (agger) was constructed across the Esquiline plateau (campus Viminalus and campus Esquillianus) providing Rome an eastern boundary (possibly) as early as the Sixth century BC (Cifani, 2013). Although funerary activities relocated beyond the agger following its construction (Albertoni, 1983; Cifani, 2013), the completion of a new circuit wall in the Fourth century BC (built in tufo di grotta oscura) triggered the accretion of a greater number and variety of distinct fringe belt land uses beyond this imposing fixation line.

The Middle and Late Republican Period: Fringe Belt Formation

Stratigraphic sequences from the area just outside the porta Esquilina show that this extramural zone maintained its pre-existing funerary character in the mid-Republic, as the presence of inhumation tombs with stone sarcophagi dating to the Third to Fourth centuries BC attest (Pinza, 1914, p. 144; Albertoni, 1983). The so-called “puticoli”—large (5 × 4 m), square tuff lined pits that housed the remains of poorer inhabitants in addition to serving as general waste receptacles—likely date to the Third century BC since they were cut into artificially deposited soils (or “scarichi”—dumps) containing materials from the Fourth to Third centuries BC (Lanciani, 1875, p. 191; Pinza, 1914, p. 165–169). The earliest private freestanding tombs, complete with interior frescos (e.g. “sepolcro di Q. Fabio”), are more or less in phase with these “puticoli” as they were constructed on the surface of these “dump” layers, likely in the later Third century BC (Pinza, 1914, p. 165–169; Coarelli, 1977, p. 207–208). Several rock-cut tombs (“a camera”) dating to the mid-Republican period were also discovered on the northeastern slope of the Caelian hill (via S. Stefano Rotondo) indicating the continuation of funerary activities further south (Santa Maria Scrinari, 1972).

Other fringe belt land uses were also present on the Esquiline in the mid-Republic, including religious institutions
and designated open spaces, such as sacred groves, sanctuaries, and temples dedicated to the funerary deities Venus Libitina and Nenia Dea (Fraioli, 2012, p. 327–328). Additionally, extramural votive deposits, likely pertaining to the lost temple of Minerva Medica, were found near the Caelian hill, north of the porta Querquetulana (Coarelli and Ricciotti, 1977). Rome’s first aqueducts, the aqua Appia (312 BC) and the anto Vetus (270 BC), also crossed the Caelian and Esquiline hills in this period, perhaps supplying water to the scared groves and/or agricultural fields located beyond the city walls (Wiseman, 1998, p. 15–16; Purcell, 2007, p. 291–294). The only literary description of what the Esquiline area might have looked like in the late Third century BC comes from Livy5 (26.10.5–6), who, writing during the Augustan age, recounts the scene of Hannibal’s aborted siege during the Second Punic War (218–201 BC):

“...the consuls ordered them to make their way through the center of the city to the Esquiline, reasoning that there would be none better suited for fighting in the defiles and among the buildings of the vegetable-gardens, the tombs and the sunken lanes running in all directions”

While this description matches well with a typical fringe belt landscape, as Purcell (2007, p. 292) suggests, this account may align better with the situation of the Second or First century BC, since the area beyond the walls underwent a series of changes at this time, likely associated with the rapid growth of the City following the Punic Wars.

In particular, paved roads were constructed on both the Esquiline (“consular via”—via Labicana) and the Caelian (“via Caelimontana”) in the Second century BC, and two more aqueducts, the aqua Marcia (140s BC) and the aqua Tepula (126 BC), were built along the Esquiline’s northern extent (Pinza, 1914, p. 167–169; Consalvi, 2009, p. 73). Additional freestanding, “singular” tombs also appear beyond the walls on the Esquiline and Caelian hills in this period (e.g., “tomba Arieti”) (Coarelli, 1977, p. 207; Consalvi, 2009, p. 73), while open spaces, sacred groves, and scattered institutional buildings persisted. For example, the guild of the tibicines (flute players) and the grove and sanctuary of Venus Libitina (associated with undertakers, funerals, and the Vinalia festival) remained in use at least until the late Republic (CIL VI 3823; Wiseman, 1998, p. 15). The (abusive) disposal of human remains and urban waste also continued in the early First century BC, as a line of cippi laid by the urban praetor L. Sentius beyond the agger (from the porta Viminalis to the porta Esquina) aimed to stop the discarding of debris, carrion, and corpses within their limits (CIL VI 31614–5; Lanciani, 1898, p. 65–67; Andreussi, 1999, p. 101). However, as Cicero6 (Pro Cluentio 37) recounts in 88 BC, the area outside the porta Esquina was not yet densely inhabited, featuring gardens and sandpits—a situation that may correspond better with Livy’s description above.

As has been shown, the evolving function of this extramural landscape fits well with the sub-phases of the fringe belt formation process as the incipient character of this fringe belt, linked to a fixation line in its fixation phase, became more pronounced as it entered into the subsequent expansion phase. However, in the Second half of First century BC a new type of land use began to appear in the area that would further impact the formation of the fringe belt on Rome’s eastern periphery, as Augustus’ companion C. Cilnius Maecenas covered one of the aforementioned pauper burial grounds for the foundation of his famed horti Maecenatis sometime after 38 BC (Hauber, 1996, p. 73). This luxurious villa-style estate was constructed in multiple stages and sprawled over both sides of the defunct city wall, extending from the Cispian Hill to the Esquiline gate (Cima, 2008).

The reclamation of the zone was famously sung about by Horace, one of Maecenas’ mentees:

“Now you can live on a healthier Esquiline and stroll on the sunny Rampart (agger), where sadly you used to gaze at a grim landscape covered with whitened bones” (Horace8, Satirae 1.8).

These so-called horti were one of several such estates constructed on the Esquiline hill that comprised numerous buildings, pavilions, and parks complete with spacious dining halls, terraces, ambulatories, pools, towers, and galleries, in addition to housing the (market) gardens from which they drew their name (Vitruvius9, De Architectura 6.5.2; Pliny10, Natural History 19.49–56; Capanna, 2012). The largest of these horti were owned by wealthy families and political personalities of the late Republic (e.g., horti Lolliani, horti Lamiani), and by the end of the First century BC these expansive estates formed somewhat of a ring or “green belt” around the old city wall, with concentrations beginning to form on the Quirnal, Pincian, Esquiline, and Caelian Hills, as well is in the Campus Martius and Trastevere (Jolivet, 1997, p. 196–197; Talamo, 2008, p. 29–33; Capanna, 2012, p. 74–78) (Figure 9). The addition of these low-density residential land uses to what had become an established fringe belt consisting predominantly of open spaces (cemeteries, groves, fields, market gardens), and institutions (temples, sanctuaries, waste removal) should be understood to signal the consolidation phase of the formation process, during which the fringe belt exhibited a dominant character.

Interestingly, the introduction of these estates to this extramural landscape may also be interpreted as a harbinger of substantial urban expansion given the predictions of Burgess, Conzen, and Alonso. Indeed, the rapidly increasing population with the situation of the Second or First century BC, especially under Caesar and Augustus, who oversaw the demolition of several residential quarters for the formation process, during which the fringe belt exhibited a dominant character.


The expansion of the CBD at the cost of housing is a prime example of land use succession and such a situation would have likely contributed to the lack of intramural space and the expansion of the continentia aedificia beyond the city walls attested to by Livy\textsuperscript{3} (44.3–5) and Dionysius of Halicarnassus (Roman Antiquities\textsuperscript{3} 4.13.5). In turn, the speed and
intensity of the urban expansion occurring would have had a substantial impact on the fringe belt that had formed on Rome’s eastern periphery.

The Imperial Period: Fringe Belt Modification

Following the formation stage, in which the fringe belt beyond the defunct city walls passed through, fixation, expansion, and consolidation phases, it would next enter the modification stage. Upon Maecenas’ death in 8 BC his famous horti were bequeathed to Augustus, and while several more such estates were added to the Esquiline in the early First century AD, many passed into Imperial possession under the Julio-Claudian emperors (e.g., horti Tauriani, Pallantiani, Torquatiani), forming a large Imperial property domain in the area (Jolivet, 1997; Talamo, 2008) (see Figure 9). The presence of these sprawling low density estates, coupled with the addition of the Castra Praetoria in AD 21 (military barracks for the emperor’s personal guard)—located just outside the porta Viminalis (and in front of the former agger)—further contributed to the consolidation of this fringe belt in the early Imperial period; however, scattered domus also began to appear east of the wall over the First century AD, indicating the further advancement of the continentia aedificia and the impending reduction of the pre-existing fringe belt there (Mols and Moormann, 1998, p. 127–130; Barbera et al., 2005; Colli et al., 2009; Consalvi, 2009, p. 78–79).

While the expansion of more intensive, residential land uses into this once peripheral zone was underway, it is difficult to assess how much it was either disrupted or stimulated by the fire of AD 64 (which impacted all but four of the 14 regions). While the Caelian hill was hit hardest by the blaze, buildings at the foot of the Esquiline were also demolished “over a vast area” to quell the conflagration (Tacitus\(^\text{11}\), Annals 15.38–40). In the aftermath of the fire many of the previously haphazard quarters that defined the Republican city were entirely rebuilt under Nero (AD 54–68) in a more consciously planned manner—e.g., no shared walls, building height restrictions, wider streets, fireproof stone rather than wooden materials, and increased water infrastructure (Tacitus\(^\text{11}\), Annals 15.43). The construction of multi-family apartment buildings, or insulae (as they are commonly but erroneously referred) also augmented under Nero (in line with public services, e.g., public baths), as these four to five story buildings provided somewhat of a solution to the growing housing crisis that had its roots in the early Julio-Claudian era (Guidobaldi, 2000, p. 140–144).

However, Nero’s decision to exploit the fire damage via the construction of his Domus Aurea (“golden house”) also drastically altered the topography of the City. This expansive estate, which connected the horti Maecenatis on the Esquiline to the imperial seat on the Palatine (domus Palatina), contained numerous ornate buildings and porticoes and featured a large lake (stagnum), open fields, thick woods, croplands, vineyards, and a multitude of wild and domestic animals—all in the (former) urban core (Suetonius, Nero 31, 39, 55; Tacitus\(^\text{11}\), Annals 15.42). Yet, despite Nero’s ambition, this building project was short-lived and the (likely uncompleted) Domus Aurea was steadily demolished after his death and the accession of Vespasian in AD 69, who returned at least part of the land to public use via the restoration of the Temple of Claudius and the construction of the Flavian Amphitheater (or Colosseum) (Suetonius\(^\text{12}\), Vespasian 9.1). As we hear from Martial\(^\text{13}\) (Liber de Spectaculis 2), significant redevelopment had indeed occurred in the Colosseum valley and on the slopes of the abutting hills by the reign of Domitian (AD 81–96):

> “Here, where the venerable mass of the far-seen Amphitheater now rises, were the ponds of Nero. Here, where we gaze with admiration at the Thermae [Baths of Titus], a boon so suddenly bestowed, a proud lawn had deprived poor wretches of their homes. Where the Claudian portico now throws its widespread shadows, was the last remnant of a felling court. Rome has been restored to herself, and what were formerly the delights of the master, are now, under thy rule, Caesar [Domitian], those of the people.”

Although the Flavian reorganization of the City gave space back to the public, many of the large civic buildings, temples, and entertainment spaces were also built at the cost of housing, forcing residential, and commercial activity into new, formerly peripheral locations (Palombi, 2016, p. 80–85), thereby contributing to the reduction of the fringe belt beyond the old city walls. As Juvenal\(^\text{14}\) (Satires 5.153–155, 6.588) tells us, the remnants of the agger, once an imposing fixation line, had become a popular common area where fortunes were told (like the Circus Maximus) and a performing monkey even entertained passersby. Similarly, Quintilian\(^\text{15}\) (Inst. Or. 12.10.74), speaks of public orations taking place in the various fora and along the agger, likely in the vicinity of the extramural Macellum Liviae and the forum Esquilinum, which by that time had also become bustling commercial zones (Coarelli, 1995, p. 298; Wiseman, 1998, p. 21–22).

In the early Second century AD, the emperor Trajan (AD 98–117) continued the “monumentalization” of the city center via the construction of a massive new forum, markets, and a monumental public bath complex—the Thermae Traiani—interring the remaining Neronian and subsequent Flavian constructions on the Oppian hill (Volpe, 2016). While such building activity would have stimulated the process of land use succession, forcing residential land uses further afield, Trajan’s continued break-up of the Imperial domain amassed by the Julio-Claudians (already begun under the Flavians) may have (inadvertently?) alleviated some of this pressure as several expansive imperially-owned horti were sold back to private owners.

---


individuals to fund the war in Dacia (Pliny the Younger16, Panegyric 50.6, 63; Talamo, 2008, p. 32). Yet, the trend of sacrificing residences closer to the center for the addition of temples and updated infrastructure continued under Hadrian (AD 117–138), and it is perhaps not surprising that the number of insulae constructed increased dramatically over the Second century AD (Guidobaldi, 2000: 146–147). By the middle of the century several of the formerly peripheral horti had also been redeveloped and subdivided into multiple domus, while domus and insulae appeared in increasing numbers outside the walls from the Porta Esquilina to the porta Caelimontana (Liverani, 1988; Mols and Moormann, 1998; Barbera et al., 2005; Cima, 2008, p. 72; Colli et al., 2009; Consalvi, 2009, p. 80).

The augmentation and densification of residences, especially on the Caelian, at greater distances from the former city walls is also notable. For example, domus from the late first and early Second centuries AD have been found under the later Lateran Basilica (Mols and Moormann, 1998, p. 123–130) while the remains of the “domus dei mosaiici” and the “domus ACEA”, both located on Via Eleniana and built in the second half of Second century AD, are directly in line with the later Aurelian Wall (Barbera, 2000, p. 105; Borgia et al., 2008). Other domus constructed during the first and Second century AD were located behind the Castra Praetoria, and therefore beyond the extent of the later Aurelian Wall, while some buildings were even immersed within it—e.g., a three-story insula next to the Aurelianic Porta Tiburtina and a Hadrianic marble officina near the Porta Asinaria (Pavolini et al., 2003, p. 85; Rea, 2010, p. 235–236; Dey, 2011, p. 79) (see Figure 10). It is worth noting here that the addition of military barracks (castra) by Trajan and Hadrian on the Caelian hill in the Lateran area is not at odds with this fringe belt modification phase since institutional land uses tend to locate within former fringe belts, especially if similar land uses are/were present there (e.g., campus Martialis) (see Juvenal 10.15–18; Whitehand, 1987, p. 84–85; Barke, 1990, p. 282–284; Colli et al., 2009). Regardless, the presence of residential, commercial, and institutional buildings this distance from the Republican city walls suggests that Rome’s continentia aedificia had greatly advanced over the Imperial period and that the pre-existing fringe belt located there had fallen victim to the processes of fringe belt reduction, which, according to the models, should signal a period of continuous urban expansion associated with concomitant economic growth.

The Late Empire: A New Fringe Belt Is Formed
Following a tumultuous period in Roman history in which the City was struck by the devastating Antonine plague (AD 165–180 and 189–190) which killed at least 150,000 inhabitants; another significant fire (AD 192); and a period of civil war following the assassination of the emperor Commodus (AD 193) the somewhat frenetic building and outward expansion that characterized the majority of the First and Second centuries AD began to wind down (Cassius Dio17 72.73; Guidobaldi, 2000, p. 152–153; Harper, 2017, p. 115). Although several new, monumental constructions were added to the City during the Severan dynasty (AD 193–235), including the Castra nova Equitium singularium on the Caelian hill, the Baths of Caracalla near the Porta Capena, and the Quirinal Temple, this period has been understood by Guidobaldi (2010, p. 318–322) to represent a phase of “musealizzazione” (or museum-izing), during which the entire city exhibits signs of maintenance and refurbishment after over a century of intense expansion and densification that resulted in the metropolis displayed on remnants of the Severan Forma Urbis.

In particular, the locations of the Castra nova Equitium singularium and the horti Spei Veteris (later, Sessorium palace) are important for investigating the formation of a secondary fringe belt beginning in the Severan period. Specifically, Septimius Severus’ decision to build a palace complex one mile east of the porta Esquilina at the southeastern extent of the City is telling. These horti Spei Veteris—named after the nearby Fifth century BC temple of Sves Vetus—were significantly augmented by Caracalla, Elagabalus, and Alexander Severus, and comprised multiple palatial pavilions, a circus, and an amphitheater, much like the Palatine palace they were designed to mimic (Historia Augusta18, Elagabalus 13.5; Barbera, 2000). Because a construction of this magnitude required ample space and a desirable peripheral site, much like the horti of the late Republic, its location could be understood to mark the edge of the now-halted continentia aedificia (Figure 10). A closer look at the location of the Castra nova also seems to indicate the cessation of outward urban growth and the fixation of a new fringe belt at this time. Firstly, its vicinity to pre-existing military installations on the Caelian hill further confirms the presence of an earlier fringe belt there; secondly, its construction at the expense of several upper-class domus (Mols and Moormann, 1998) suggests a building trough and a general “housing slump,” likely fueled by lack of residential demand—a situation further evidenced by the diminishing construction of insulae, which eventually came to an abrupt stop by the later Severan age (Guidobaldi, 2000, p. 153; Guidobaldi, 2010, p. 322).

Following the assassination of Alexander Severus in AD 235, the conditions for the continued formation of this fringe belt were especially prevalent, as the next 36 years saw 60 attempts to capture the Imperial seat, resulting in a period of political disarray that contributed to a lack of public building and imperial investment at Rome (White, 2004, p. 27–31). In fact, by the mid-Third century AD many of the pre-existing multi story, multi-family apartment buildings (insulae) were converted into (often luxurious) single-family domus—the direct opposite to what was occurring the previous century (Guidobaldi, 2000, p. 152–154; Guidobaldi, 2010, p. 322). From AD 249-270 the Plague of Cyprian swept through the City further damaging a population that probably never fully recovered from the prior Antonine

18Historia Augusta (Translated by D. Magie 1921). London: Harvard University Press.
Mandich Ancient City, Universal Growth?

FIGURE 10 | Map of Rome showing Republican and Aurelianic Walls. NB: Location of main roads; Porta Asinaria and Porta Tiburtina; the Castra Pretoria; and the Amphitheatrum Castrense (https://commons.wikimedia.org/wiki/File:Plan_Rome-_Aureliaanse_Muur.png) (Joris1919) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0/)].

plague (Harper, 2017, p. 136–138), and, apart from the Baths of Decius (built on the Aventine in AD 252), very few large public buildings or monuments were added to the cityscape between the Severan dynasty and the reign of Aurelian (Guidobaldi, 2010, p. 322–323).

The construction of the Aurelian Wall, begun in AD 271, can then be seen to formerly conclude the previously unchecked urban expansion of the City; however, an examination of its course and composition can shed further light on the formation of a fringe belt during the Third century AD. To begin, roughly one sixth of the Aurelianic circuit is comprised of pre-existing structures and many more were demolished or buried to make way for both an external ditch (fossa) and an internal access road (intervalum) (Lanciani, 1892, p. 106; Dey, 2011, p. 165–166). The variety of sites revealed by excavations around the wall is striking, as both private and imperially owned (and procured) properties (often referred to as villas or horti in Nineteenth century documentation) were impacted by its construction (Lanciani, 1892, p. 104–110; Richmond, 1930, p. 11–16). Starting in the north, several estates from the mid-Imperial period near the Castra Praetoria were dissected by the wall and then abandoned, as were other domus built in the first to Second century AD located beyond the castra, past the Wall’s extent (Lanciani, 1892, p. 104–106; Pavolini et al., 2003, p. 85; Guidobaldi, 2010, p. 316). On the east, a decorative garden nymphaeum complete with colored pumice, shells, and statuary niches (with intact statues) was absorbed by the circuit near the porta Tiburtina along with several larger tombs in the same area (Lanciani, 1892, p. 104, 109). Moving south, the recent discovery of a quarry and small farm near the porta...
Asinaria also attest to characteristic fringe belt activities in the area (Rea, 2010, p. 232–238).

Given this information, it appears that a pre-existing fringe belt was exploited for the construction of the Wall since its course attempted, where possible, to follow a path of least resistance, cutting through peripheral estates, funerary zones, and less densely built plots to more quickly (but roughly) fortify the City on its eastern side. Although the course of the Aurelian circuit was influenced by the presence of the tax border, pomarium, and other practical, geographic, and militaristic considerations (see Palmer, 1980; Coarelli, 1997; Dey, 2011, p. 72–86), its relatively star-shaped perimeter encompassed areas of significant urban expansion that followed major thoroughfares and aqueducts beyond the Republican city walls, matching well with the predictions of bid rent theory and highlighting the dynamic physical expansion of the City over the period in question (Mandich, 2015, p. 85–92). Yet, because the course of the Wall seems to have cut through an established fringe belt, leaving certain quarters outside its limits, the circuit should not be viewed as the quintessential maximum extent of ancient Rome, but rather as a sort of “cookie-cutter” that separated the most densely urbanized areas from the more extensive urban development that had spread uninhibited into the immediate countryside—once giving Dionysius the impression of “a city stretching out indefinitely.” While the construction of the Aurelian Wall can be understood to mark a period of contraction and consolidation, it also gave the City a clear, physical urban extent, providing a new fixation line that served to restart many of the complex morphological processes that had slowly played out over the previous six and a half centuries.

A Glance to the Suburbs: Fringe Belt Translation

To more fully understand the processes of urban expansion and economic development occurring in ancient Rome, a brief look at sites in its immediate surroundings is required. Although mid-Republican villas are often elusive in the archaeological record, evidence from the suburbs does attest to a network of productive villas surrounding the City and taking advantage of the economic situation created by its growth and vicinity (De Neeve, 1984; Carandini, 1985; Morley, 1996, p. 55–63). In particular, properties on the Centocelle plain (4th mile of the via Labicana) and in the eastern environs were engaged in agricultural production (often viticulture and poly-culture) and consisted of farmhouses (typically built in opus quadratum and/or opus incertum) surrounded by fields and demarcated by ditches, banks, and/or roads (fossae or rivi finales), each holding about 50–70 iugera (12–18 hectares) (Musco and Zaccagni, 1985, p. 91–106; Volpe, 2004, p. 448–455; Volpe, 2008, p. 262–263). According to Volpe (2000, p. 186), the construction of these mid-Republican villas represents a period of change in which a previous, more autonomous, domestic system mutated, as villas began to reorient around Rome due to the increasing urban demand for agricultural goods (likely in accordance with the predictions of Von Thünen’s model).

In the later Republic, many of these earlier villas were augmented (or obliterated and rebuilt, frequently in opus reticulatum) via the addition of more luxurious residential quarters, porticos, and ornamental garden spaces—often at the expense of areas previously dedicated to agriculture (Volpe, 2008, p. 262–267; Volpe, 2009, p. 381). In some cases this could signal the reinvestment of profits by villa owners engaged in successful agricultural practices; however, it seems more likely that these properties now belonged to a new class of owner as the value of such estates would have increased considerably in line with the physical and demographic growth of the City (see Figure 5). In fact, many more villas were constructed in this period overall as 71 of the 100 villas considered by De Franceschini (2005, p. 297) in her book Ville Dell’agro Romano were built ex novo between the Second century BC and the Augustan era, with 50 constructed specifically between the later First century BC and the early First century AD—indicating an influx of people and wealth into the countryside that coincided with a period of rapid urban expansion.

Over the First century AD the presence of imported wine amphorae in both urban and suburban contexts also escalated dramatically (Volpe, 2009, p. 380–381). While this is often linked to a variety of factors, including a growing urban population and a shifting market, it also reflects a more densely settled hinterland that had shifted in the late Republic from extensive agriculture (such as viticulture/poly-culture) to intensive practices, primarily for the urban markets—e.g., horticulture (fruits, vegetables, herbs), pastio vilactica (rearing of small animals and birds), and the cultivation of flowers and plants for garlands and dyes (in hortensia) (Carandini, 1985; Patterson, 1987; Witcher, 2006; Volpe, 2009). Although vines were probably still widely cultivated on suburban properties during the Imperial period, wine was likely produced in smaller quantities of higher quality (e.g., the famed vite Nomentana) and transported/stored in barrels, making it archaeologically undetectable, explaining both the absence of Italian amphorae and the augmentation of provincial imports (De Sena, 2005, p. 6; Volpe, 2009, p. 280–281).

Such a situation would suggest that land in Rome's immediate suburbs had become more expensive due to increasing demand, in turn, leading to a denser landscape of monumental villas that included more expansive residential quarters, ornate interior décor, large cisterns, and grandiose bath complexes (De Franceschini, 2005, p. 304–314; Volpe, 2007, p. 395–398). Indeed, by the mid-Imperial period villas on the Centocelle plain were now situated c. 700 m from each other with each one likely belonging to a different owner (Volpe, 2000, p. 204) (Figure 11). As Coarelli (1986, p. 54) has suggested, by this time these properties had probably come to function as primary residences (or domus) for their owners given the nearby viae and the outward expansion of the City that provided them a more peripheral location. In fact, the presence of high-end villa estates near major roads and aqueducts matches well with the predictions of bid-rent theory since infrastructural amenities work to steepen residential bid rent curves in their vicinity, making such properties more expensive (Alonso, 1964, p. 110, 141–142) (Figure 6). As such, it is perhaps not surprising that many estates in the eastern suburbs were either owned by
senators, often of provincial origin, or connected to members of the Imperial family (see Pliny the Younger\textsuperscript{19}, \textit{Ep.} 6.9.1–6; Coarelli, 1986, p. 41–55).

While these estates still maintained profitable agricultural components, the architectural and functional changes observed should be directly linked to Rome's urban expansion and the processes of land use succession and fringe belt reduction and translation. Since the former fringe belt on the Esquiline and Caelian Hills was heavily reduced by a steadily advancing \textit{continentia aedificia} over the First and Second centuries AD, certain land uses may have translated further afield (see Figure 8). Specifically, previously peripheral low-density estates (such as the \textit{horti}) would have been forced to migrate to a new peripheral location, while the conversion of many \textit{domus} into \textit{insulae} over the course of the Second century AD could be related to the relocation of these single-family home owners to more removed suburban estates (Jolivet, 1997, p. 201–203; Guidobaldi, 2010, p. 321–322).

Because the translation of a fringe belt is often associated with the presence of another further removed fixation line, the conversion of primarily productive properties into more luxurious residential estates around the fourth to sixth mile markers of Rome's major \textit{viae} is of particular interest since this location marked the extent of Rome's fabled \textit{ager Romanus antiquus}. Although debates concerning the antiquity and existence of this enigmatic perimeter continue, this distance did hold a juridical, practical, and memorial significance into the Imperial period (see Strabo\textsuperscript{20}, \textit{Geography} 5.3.2; Appian\textsuperscript{21}, \textit{Civil Wars} 1.23, 1.57; Smith, 2017), possibly contributing to a perceived discontinuity in the countryside that provided a potential fixation line around which such residences could conglomerate.

Turning to the Third century AD, despite Rome's damaged population elite competition remained strong in the environs. The construction of an aqueduct, widely attributed to Alexander Severus, running rather irregularly (and clearly off the path of previous aqueducts) through this eastern zone should signal the continued importance of properties there, many of which were now owned by the Imperial family, either through inheritance, or confiscation (Coarelli, 1986, p. 51, 56–58). Examining the sites of the Centocelle plain, the so-called “\textit{villa delle terme}” shows evidence for significant investment into the bath complex in the Second and Thrid centuries AD, while the so-called “\textit{villa delle piscine}” exhibits a similar pattern of upgrades (Coletti, 2007, p. 201–213; Volpe, 2007, p. 399–400). In the area of Torre Spaccata (2 km east of Centocelle), “villa A 204” was redeveloped in the early Third century AD, as two \textit{arae} for grain macination and a \textit{torcular} (wine/oil press) were demolished and paved over with marble \textit{opus sectile} floors, indicating the expansion of the \textit{pars urbana} at the expense of the \textit{pars rustica} (Ciceroni, 2008, p. 211–214). Although elite investment continued at certain pre-existing sites over the Third century AD, as is expected during an economic “trough,” almost no new villas were constructed in Rome's suburbs at this time and the number of villas abandoned rose considerably, especially following the construction of the Aurelian Wall (De Franceschini, 2005, p. 297).

\section*{DISCUSSION}

A re-examination of the available evidence from ancient Rome's eastern periphery has shown that the processes of fringe belt formation, modification, and translation were unfolding in ancient Rome much as they would in a contemporary urban setting. The increasing number of distinct fringe belt land uses appearing beyond the defunct city walls over the Republican period points to the formation of a fringe belt that experienced fixation, expansion, and consolidation phases. The subsequent expansion of public, civic, and commercial land uses in the city center during the Imperial period would have then forced the relocation of certain land use activities, and the changing function of residential plots on the City’s expanding periphery—from \textit{horti} to \textit{domus} to \textit{insulae}—matches well with the predictions of bid-rent theory and the process of land use succession. This could suggest that the outward expansion of the City was linked to a combination of increasing population, innovations in transport technology, and (possibly) rising per capita incomes, which, when combined, should represent phases of substantial economic growth. As the city continued to expand outward over the Imperial


\textsuperscript{20}Strabo (Translated by H. W. Jones 1917–32) \textit{Geography}. London: Harvard University Press.

\textsuperscript{21}Appian (Translated by H. White 1912–13) \textit{Civil Wars}. London: Harvard University Press.
period many primarily productive villas were transformed into upper-class residential estates as land values would have elevated, especially near roads and aqueducts where bid rent curves steepened. However, following the Antonine plague, and a period of political turmoil, the advancement of the continentia edificia slowed dramatically as a new fringe belt began to form on the edge of halted urban periphery, reflecting a building slump associated with a period of overall economic stagnation.

Using the theoretical framework of settlement scaling theory, it appears that the same correlations found between population, infrastructure, area, and socio-economic outputs in modern cities also appear to have been at play in ancient Rome, although this has yet to be confirmed empirically. Apart from the archaeological evidence presented above, the ancient sources also suggest that densification was occurring over the period of study. For example, Livy\(^3\) (21.62) mentions that multi-story buildings already existed in 218 BC, while more general accounts of Rome’s Republican urban fabric speak to a jumbled maze of streets and (often poorly built) high-rise structures (Livy\(^3\) 5.55, 40.5; Tacitus\(^11\), Annals 15.38–43). In the Mid-First century BC the construction of multi-story residences increased (Vitruvius\(^3\), De Architectura 2.8.17), and following the Augustan reforms, and a period of peace and stability ushered in by the new monarchy, the building, sale, and remodeling of houses apparently became “unceasing” (Strabo, Geography\(^20\) 5.3.235). The proliferation of insulae over the First to Second centuries AD also points to continued densification in the Imperial period, and despite attempts to maintain a more controlled brand of urban development following the fire of AD 64, 10 story buildings remained prevalent (Juvenal\(^14\) Satires 3.407-8). In addition, data from modern cities has shown that for each doubling of the population, land rents (i.e., bid-rent curves) rise by 50% (Bettencourt, 2013b, p. 6)—a statistic that may lend some doubling of the population, land rents (i.e., bid-rent curves) rise.

In sum, Rome’s expansion fits well with the principles and predictions of bid rent theory, the fringe belt model, and settlement scaling theory. The successful application of these models has allowed to view the physical growth of the city as a new dataset to examine economic development in the Roman period. Furthermore, if the same or similar processes driving the physical growth of contemporary cities were occurring in ancient Rome, it can be argued that ancient Rome was growing and evolving much like a modern metropolis. If correct, this allows for the urban growth of ancient and modern settlements to be linked regardless of temporal or technological differentiations, illustrating the continuity of human settlement dynamics, and providing different ways to analyze and interpret urban growth, decline, and economic development in the ancient world (see Mandich, forthcoming).

**AUTHOR CONTRIBUTIONS**

The research and writing for this paper was all undertaken by MM.

**FUNDING**

This paper was funded via an OPEN-AIRE grant to Dr. Francesca Fulminante through the University of Rome (Marie Curie Sklodowska Fellowship 628818- Past-People-Nets 2014–2016).

**ACKNOWLEDGMENTS**

This paper was published thanks to an OPEN-AIRE fellowship granted to Francesca Fulminante for the Marie Slodowska Curie Project 628818 Past-People-Nets. Special thanks are owed to Francesca Fulminante and the other members of the editorial committee for inviting me to contribute to this Special Issue and for funding the publication of this research. Many thanks are also owed to the Reviewers for their constructive comments.

**REFERENCES**


Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Mandich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Urbanity as a Process and the Role of Relative Network Properties—A Case Study From the Early Iron Age

Oliver Nakoinz*, Michael Bilger and David Matzig
Institut für Ur- und Frühgeschichte, Christian-Albrechts-Universität, Kiel, Germany

In the past numerous concepts of urbanity have been discussed and a variety of criteria for towns have been developed. They include size, population, legal aspects, way of life, structural and functional approaches. However, since the mentioned criteria cover only a part of the phenomenon and partly use fixed and arbitrary thresholds, they are not sufficient for analysis. We turn to an understanding of urbanity as a process that creates and shapes the scenery of the buildings and people and that is mainly driven by complexity. In this sense, we understand urbanity as a process of adaptation to changing conditions or contexts in a complex settlement system, which is triggered by size, attracted by exemplary solutions and characterized by the emergence of new structures. In this paper we address the issue of relative centrality as proposed by Christaller in the urbanity process as well as centrality within a network sciences approach. Our aim is to interweave different concepts of urbanity, centrality, interaction and connectivity, combining different concepts and research traditions as well as expanding them, resulting in a collection of different terminological frameworks. In the context of adaptation, urbanity is relative in the sense that different places may have gained better or worse adaptation under different conditions. The urbanity process is always shaped by the threat of too much complexity and too little connectivity. Above all, it is a certain surplus of connectivity that characterizes urbanity. This surplus is mapped by the variant of centrality proposed by Christaller. While Christaller’s models can be transferred into network sciences frameworks, Christaller does not offer an adequate centrality measure. Therefore, his concept of centrality cannot be transferred correctly without being translated carefully into the network research context. In this article, we argue why this is necessary and explain how it can be done. In this paper the above concept will be applied to the Early Iron Age Princely Seats with a special focus on the Heuneburg. In order to represent similarities and interaction between different nodes a very limited part of the material culture can be used. For this purpose we use fibulae which allow for fairly accurate dating and hence ensure a narrow time slice for the network analysis. Using Fibulae the research will be limited to a certain social segment, which we refer to as “middle class.” This paper is intended to deal with the rather complex issue of urbanity using more simple approaches such as network analysis. In this context, we pursue a tight integration of theory and methodology and we consider certain conceptual
1. INTRODUCTION

One of the reasons for the emergence of cities is the structural advantage of central places. Interpreting towns as central places allows us to explain the location of many towns and pre-urban settlements. Despite the fact that there are numerous publications on urbanity and centrality, the theoretical foundation for the combination of these two main concepts appears to be rather weak. This paper attempts to explore the relationship between the two concepts as well as between the two and other concepts. The keywords which will be addressed in the theoretical part of this paper are urbanity and centrality, the aforementioned main concepts, as well as complexity, connectivity, and interaction. We show that the main concepts are closely linked, which reveals new facets regarding the other keywords. Theoretical concepts have a significant impact on methodology. According to our understanding of urbanity and centrality in the light of complexity and connectivity, some changes in the methodology for the estimation of centrality in geographical networks are required. These changes have been directly applied in a case study on Early Iron Age and concern the concepts for the integration of network centrality (cf. Taylor et al., 2010) and Christaller centrality (cf. Christaller, 1933). Since the methodology is now completely covered by theory, these changes lead to a substantial surplus in the interpretations.

This paper mainly focuses on the so-called Early Iron Age Princely Seats (Kimmig, 1969; Krausse, 2008) which are central to the discourse on early urbanization. The Princely Seats are characterized, among other things, by Mediterranean imports, fortifications and rich graves. Some scholars consider these sites to be the ‘first towns north of the Alps’ (Krausse et al., 2016) while others do not use as many exciting superlatives and buzzwords when referring to them, but focus more on complicated relationships and processes (Stoddart, 2017). Perhaps the question whether princely seats are towns or not is not of any greater importance. Regarding the knowledge of both the social and economic circumstances at that time the question which processes took place in the context of what we call early urbanization is more crucial. This paper aims to contribute to the clarification of this question based on research done in South-West Germany (Figure 1). Our main objective is to explore the role of people in urbanity processes who were not part of the power elite. We decided to use fibulae as an artifact type, as they play a significant role in terms of chronology and in some way are related to the social fraction that we refer to as middle class due to the absence of a more suitable term. Although, fibulae also appear in elite graves, the majority of the fibulae can be attributed to moderate social ranks.

2. THEORETICAL BACKGROUND

In this section we take a closer look at urbanity and then go into more detail with the discussion of some keywords such as complexity and centrality. In this context, we apply a hermeneutic approach starting with the investigation of certain aspects of the respective terms and looking at their relationship, which is followed by investigating the next set of aspects building on top of the already made relationships. Using this approach we may not be able to avoid redundancies, but it seems to be the most suitable one to clarify the relationship of the tightly connected concepts.

2.1. Concepts of Urbanity

In the literature of different disciplines numerous concepts of urbanity are known. Covering all of them would exceed the scope of this paper. However, some of the concepts need to be addressed (Nakoinz, 2017b). First, we would like to mention the large number of quantitative approaches including legal issues and lifestyle aspects. Towns can be defined as places which have a so-called town law. The town law regulates the behavior of the people who live in that town or toward the surroundings. The medieval town law of Schleswig (Hasse, 1880) is an example which shows that in towns many aspects of daily life require proper regulation, which is not the case for villages. That difference is important, even though the required written sources are not available for prehistory. By introducing laws, rules become institutional and it is obviously the existence of these institutions which is important for towns. Another important qualitative aspect is the specific lifestyle of people living in towns (Wirth, 1938). One example is the increasing anonymity in towns, which is characterized by a lack of economic and to some extend even social interaction between individuals. Two further aspects worth to be mentioned are specialization and social mobility. The dynamic life in towns is another important aspect which, however, is difficult to specify. Our idea of a proper town is a place in which every day comes with some new ideas, events, information and opportunities, rather than a place in which everyday life is characterized by uniformity. While some qualitative aspects are clearly related to the size and complexity of a place, elements of a urban lifestyle can also be found in rural settlements (Gans, 1962). Therefore, many qualitative aspects can be described as symptoms of urbanity, rather than being the main or even diagnostic features of urbanity.

Now, we can turn to the more simple quantitative approaches. Eurostat for instance (Eurostat, 2018) uses a threshold of 5,000 inhabitants to define places as towns. However, this approach appears to be problematic, because that way the boundaries of a place could simply be moved further and further to reach...
a certain threshold. Aggregating several villages to a unit of 5,000 inhabitants is certainly not the same as a densely occupied place with 5,000 inhabitants. Density defined as population by size can be used to overcome some of these problems. Eurostat (2018) defines a density threshold of 300 people per square kilometre (ppskm) and Demographia (2015) uses a value of 400 ppskm. This approach, however, is still not convincing, since the thresholds are arbitrarily determined and hence the classification is rendered meaningless. Instead of such arbitrary numbers, Roland Flechter’s calculations on settlement growth could be of help (Fletcher, 1995). It has been shown that official population thresholds in different countries range between 200 and 50,000 inhabitants (Deuskar, 2015), which makes it impossible to compare cities around the world. A more convincing and popular group of approaches is based on economic functions (Smith, 1989). Although archaeological indicators are even more problematic, this functional approach provides a better insight into the mechanisms of urbanism. Towns have less agriculture and a larger craft and service sector than rural villages. Furthermore, they show a certain degree of division of labor. On the one hand, a specialization takes place which leads to a local division of labor. This means that each craftsman is responsible for a different step in the chaîne opératoire and delivers the semi-finished product to the next expert. On the other hand, there is some kind of geographical division of labor. Towns appear to be consumer sites regarding agricultural products, while producer sites can be seen as consumer sites using craft ware. The division of labor eventually leads to more interaction with other people as well as an increasing dependency on them.

The division of labor leads directly to the structural approaches that focus on the relationship between places and people. Centrality concepts are the most important instance of structural approaches. The term central place was developed by Christaller (1933), in order to circumvent the complex discussion of urbanity. However, the term itself turned out to be complex being loaded with baggage from different disciplines, such as legal and historical approaches. Christaller’s idea was to use formal characteristics to explain why places in certain locations have a certain size. He found that the size of a place depends, among other things, on its relationship to other places. Therefore, he isolated this relational aspect as centrality. According to Christaller, centrality is defined as the relative importance of a place (we prefer to translate “Bedeutung” rather with “importance” than “meaning.” However, size and importance of a place are not identical. Christaller defined a central place as a place that has a certain surplus of importance compared to the importance it would have if it had the same size, but no connection to other places. Obviously the idea behind this is that there might be places with many people. However, if these people did not interact with people from other places, the said place would be of no relevance. Christaller hence defined centrality as mentioned above. The importance in this case is provided by central functions such as production, administration and trade with other places. From an urban point of view, centrality is an important aspect of urbanity. Nevertheless, it does not cover all aspects of urbanity. Even small and isolated places can have a certain degree of urban lifestyle. However, we assume that in general towns can be considered as central places. In addition to the definition of centrality and an empirical case study, Christaller presented different models that demonstrate how places obtain a certain degree of centrality. The models are straightforward and based on the assumption of transport cost minimization. This means that people try to minimize their transport costs and hence go for the nearest possible place for supplies. It results in the emergence of well-distinguished,
can be supplied.

There are further aspects of Christaller's centrality model and a lot of literature on central place theory. For more details refer to Knitter and Nakoinz (2018) and Nakoinz (2019). Instead of going too much into detail now, we turn to another concept of centrality, which has been derived from social network theory (Freeman, 1978). This type of centrality refers to the strategic location of a place within a network, rather than the place gaining importance by supplying certain areas with goods and services. In this context, a network center for example plays a significant role, as it guides the interaction between different places.

By taking a little step back we might observe that both approaches look at different organizational structures. Since every place has a certain degree of both Christaller centrality and network centrality, these two organizational structures are complementary. A combination of the two approaches seemed appropriate and was developed during the previous decade (Nakoinz, 2012, 2013, 2019; Knitter and Nakoinz, 2018). First of all, we should follow the two traditional approaches and define centrality as a relative concentration of interactions. Subsequently, we are able to define different models including the Christaller model and the network model. For more details please refer to Nakoinz (2019), where an extensive discussion of the topic is provided. In this paper we will focus on a specific aspect of the integrated approach, which is highly relevant for the urbanity discourse and which has important methodological implications. First, however, we make a small excursus looking at the different paradigms involved.

In archaeology, network approaches are supposed to replace the old paradigm of cultural, ethnic, and other groups. They have their focus explicitly on relationships and therefore are closely intertwined with modern research topics. And although they have entered the archaeological world only recently, they have a long history of research and two completely different and independent roots. In this context, we first need to mention geographical or spatial networks with a research tradition going back to the 1930s and it experienced great success at the same time as the geographical network tradition (Scott, 2000; Freeman, 2004). The said period is called "Harvard Break Through" (Raab, 2010). Both traditions are part of the quantitative revolution of the 1960s. Today, both traditions are under discussion, although mainly the social networks have been in the spotlight of current research. There are even attempts to combine the two approaches (cf. Knappett et al., 2008; Rivers et al., 2013; Barthelemy, 2014; Prignano et al., 2019), which would definitely be the right step.

In the previous one or two decades network analysis became an important topic in archaeology (Knapett et al., 2008; Brughmans, 2010; Collar et al., 2015). On the one hand, archaeologists have been trying to focus on the more modern social network approaches. However, leaving geographical network approaches behind turns out to be difficult, since in archaeology spatial information is an essential aspect. Case studies have shown that archaeological researchers might develop a network approach for their own field which could serve as integrative concept for other disciplines as well. The archaeological application of network analyses solves archaeological problems by using ideas from both the geographical and the social network tradition (Sindbæk, 2007; Filet, 2017; Wehner, 2020).

In recent years the status of archaeological network research has been discussed (Fulminante, 2014). The main goal has been to close the theory gap in network analysis by developing an archaeology of interaction (Knappett, 2014).

We would like to go one step further and combine grouping approaches with social and geographical network approaches to create complementary approaches and present one coherent concept (Nakoinz, 2017a). The different approaches are closely linked, but are subject to different traditions and terminological cultures. This article will show how we can benefit from the combination of different so-called paradigms. Furthermore, it will become clear how important the careful use of terminology is.

After demanding integrative approaches, a combination of network and clustering approaches, we would like to come back to the issue of urbanity and hope that the reader of this paper will keep the “integrative spirit” from the previous sentences in mind. The problem with the mentioned concepts of urbanity is that although each concept covers a specific aspect that is important, there is no concept which covers all of the facets. Furthermore, the essence of urbanity is not captured at all. Towns are characterized by a certain size on a relative scale. However, it is not the size itself which defines a place as urban. It is rather a certain reaction to the requirements of size including coping with the emerging complexity and benefiting from the opportunities of big sized places (Nakoinz, 2017b).

Complexity is a concept that can be derived from a certain degree of interaction and connectivity and hence needs to play a major role in the urbanity discourse. In the following section we are going to have a closer look at those keywords.

### 2.2. Complexity, Interaction, and Connectivity

The meaning of complexity has many facets. For our approach we use the concept of complexity that stems from complex system sciences (Gell-Mann, 1995) and was developed on top of system theory (von Bertalanffy, 1968), as we found it to serve our purpose best.

A system is a model of a specific research topic that maps not only the structure, but also the dynamic interrelationships between different elements. A complex system is a system that shows non-linear behavior, path dependency, butterfly effect and some other features and which is not predictable. Complexity in this sense can be caused by the existence of too many elements or by the elements being too complicated, especially if these elements are in certain non-linear interrelationships. Complex networks can involve some remarkable effects such as
emergence, butterfly effects, and path dependency. Cities can be considered as complex systems (Jacobs, 1992; Batty, 2005; Bettencourt and West, 2010; Bettencourt et al., 2013; Schläpfer et al., 2014; Nakoinz, 2017b), even though they are different from complex biological systems (Bettencourt, 2013). Complexity-based urbanity research is a field which aims at combining social and geographical networks by developing social networks of settlements embedded in a geographical space (Ortman et al., 2014).

In addition to the complex nature of social interaction, two main drivers of complexity are size and population. It is a fact that human beings can only manage a certain number of effective contacts (Dunbar, 1992). In this paper, effective contacts are defined as interactions between people that usually have a social, economic or cultural impact. There are thresholds for the size of communities (MacSweeney, 2004; Feinmann, 2011) that define at what point these communities become unstable and thus obtain a certain degree of instability.

Communities that exceed the threshold and become unstable can regain a certain degree of stability by applying complexity reduction methods. This can be achieved by restricting the number of effective contacts, for example within-group connections, neighbors or the edges of a hierarchical network of individuals. The limitation of effective contacts within hierarchies is the most efficient method of complexity reduction. This does not necessarily mean that people from different groups and hierarchies are not allowed to talk to each other. However, there are certain “cultural rules” that make these contacts less effective.

Let us take a closer look at what causes complexity. In fact, it is not the size of a community, but rather the number of effective interactions that causes the threshold to be exceeded. A high number of interactions can also lead to people being more stressed and making mistakes as well as something we could call mismanagement of connections. The higher the number of effective interactions, the more unpredictable is the urban environment to the individual and the urban system to all people. In this sense, interaction can be identified as a distinctive feature of complexity, since in a predictable system the number of manageable connections is limited. Another aspect of complexity is connectivity, which plays an important role in terms of making decisions. If there is the opportunity for people to interact with each other, these people first need to decide whether they actually want to interact with each other or not. In short, when it comes to connectivity, we use our cognitive capacities to decide who should and who should not be our interaction partner and how to manage that contact. When it comes to interaction, we use our cognitive capacities to make decisions regarding the content of an interaction. However, this puts additional stress on people resulting in further uncertainties, especially in systems that are already hard to predict. It is obvious that interaction and connectivity are somehow related. We decided to include these terms in our paper, as they seem to be essential to our topic. In the following section we will present a definition for both terms.

Interaction is a process of action that involves at least two interacting partners. In this process the two interacting partners can have a common goal. In that case the interaction may lead to an expansion of each partner’s capabilities. However, if the interacting partners have contradicting goals, this may lead to a reduction of each partner’s capabilities (Arponen et al., 2016). While interaction is a real action, connectivity refers to possible actions and thus can be defined as potential interaction. However, the aspect of potential action requires a more detailed discussion. For now, defining connectivity as potential interaction is sufficient and we can conclude that both interaction and connectivity are characteristic features of complexity. The difference between interaction and connectivity is that connectivity involves connections that are not used for actual interaction. Maintaining connectivity requires some effort. However, it does not come along with taking advantage of any benefits. It is the interaction through which one is able to reap those benefits. Thus, performing an interaction requires additional efforts.

Later on, we will discuss the effects which connectivity has on complexity in an urban context. First, however, it should be noted that a higher degree of connectivity and thus complexity go along with more opportunities. Many potential connections raise the chance for a division of labor, as it becomes more likely to find someone who participates in production by conducting a step in the chaines opératoires. In addition, a division of labor increases connectivity by maintaining the social connection between collaborators. It also leads to more interactivity, as the collaborators are required to interact. It becomes apparent that connectivity and a division of labor mutually reinforce each other. A high degree of connectivity increases the chance of more knowledge to be exchanged and hence the chances for innovation. Finally, a high degree of connectivity increases the chance to offer central functions. At the same time, centrality leads to more connectivity and actual interaction at central places. While centrality also leads to more complexity at central places, it reduces complexity in non-central places. This is due to the fact that in central settlement systems fewer decisions have to be made. In other words, centrality leads to a shift of complexity from the periphery to the center.

Increasing the degree of connectivity and thus of complexity means increasing stress. As mentioned above, a high degree of connectivity requires a large number of decisions and leads to scalar stress (Alberti, 2014). Numerous connections have to be maintained and if the number of connections exceeds a certain threshold, the system becomes unpredictable.

Increasing the degree of connectivity and thus of complexity also means a shift in lifestyle. In this connection, three different aspects should be mentioned. If the number of connections exceeds the already mentioned threshold, the system not only becomes unpredictable, but people start to drop some of their connections due to their limited capabilities. Dropping or devaluing social contacts leads to nothing but an anonymous way of life, which has already been named a characteristic feature of urban life. As a consequence, rules will be introduced to fight the unpredictability of complex urban systems. For this purpose, institutions are being established. Institutions are characterized by specific social functions and roles, where the role incumbent might change, whereas the function persists. That way, it can be ensured that the functions will always be fulfilled in the same way, which leads to a higher degree of stability and predictability and thus reduced complexity.
In summary, agglomerations constantly tend to become too complex and thus too difficult for their inhabitants to live in. This is particularly true for central places with a high complexity, as opposed to central places of the same size that do not have the same level of complexity. Settlement systems with several central places are characterized by rather low centrality in the space between the central places and where each central place shows a rather high degree of complexity. In the next section we will explore which role centrality plays for urbanity.

2.3. Urbanity and Centrality
Prepared with our knowledge regarding complexity and connectivity, we can now come back to the characteristic features of urbanity. It becomes apparent that places with a variety of opportunities trigger and drive different urban processes and towns appear to be places in which new opportunities arise due to urban processes. This mutual interrelation roots in a certain degree of complexity and in turn leads to more complexity. It is the aspect of connectivity in particular that produces both new opportunities and additional complexity. For the following part, we can state that the actual degree of interaction at urban sites is higher than it would be at non-urban sites of the same size. Here, we reach the point where the relative surplus of importance of Christaller’s concept is inevitably for the urbanity discourse.

These characteristics do not provide us with fixed parameters for the description of a place. However, they appear to be related to specific urban processes as described above. Therefore, urbanity should be understood as a process rather than a status or property of a place. It can be defined as the search for opportunities and the attempt to cope with unpredictable and highly connected environments. In fact, the term ‘urban jungle’ derives its meaning from an unpredictable and dense environment in which unforeseen interactions can interfere with the predicted course of life.

Not only do the above described processes involve a high degree of connectivity and interaction, but the existing connectivity also continuously increases. Centrality, which we consider as a concentration of interaction (Nakoinz, 2019), is highly involved in this process. While centrality leads to a reduction of complexity at the periphery, it further increases it at the center. We already described the two basic concepts of centrality above. Now we need to take a more detailed look at what centrality actually means and how it is measured. First, we discuss Christaller centrality. While our definition of centrality as the relative concentration of interaction is very general and covers both Christaller’s approach and the network approach, Christaller defines centrality as the relative surplus of importance of a place. In order to understand the significance of Christaller’s concept of urbanity, we need to clarify the meaning of “relative.” Since this aspect of Christaller’s central place theory sometimes is neglected and is partly not covered in the secondary literature, we feel that it needs to be explicitly included. A considerable amount of literature on central places exclusively focusses on an absolute importance instead of a relative surplus of importance and hence does not address centrality as defined by Christaller at all. In this paper, we adopt Christaller’s ideas and describe them using our own terminological and conceptual frame. For the original text and concept please refer to Christaller (1933).

We assume that there are different attractors of interaction at a place. First-order attractors, which are analogous to first-order effects in point pattern analyses, are based on locational factors such as soil, water, and natural resources. These factors determine the carrying capacity at a place. Second-order attractors attract interactions by structural properties such as network integration. Both types of attractors represent an interaction potential we call primary interaction potential. Basically, primary interaction potentials refer to the attraction of people. Secondary interaction potentials are based on the primary ones and represent the number of possible interactions. They refer to connectivity, which can be defined as potential interaction. People are attracted by environmental benefits such as a high carrying capacity or natural resources (primary centrality potentials) and by the opportunity to make new useful connections (secondary centrality potentials). The latter is influenced by the number of people at one place, the difficulty of making new contacts and the feasibility to make contacts with people from other places. Since there is no linear relationship between primary and secondary interaction potentials, it makes sense to distinguish between the two potentials. It is the complexity of social relationships that makes it impossible to use the primary potential as proxy for connectivity, even though these two aspects are closely related. The secondary interaction potential already takes these aspects into consideration. In order to circumvent the complicated deduction of the secondary interaction potential based on the first one, structural properties are being used as proxy for the interaction potential. In the case of Christaller centrality, the number of people represents the probable number of interactions. In case of network centrality, the probable amount of interaction will be predicted based on network properties. This information can be provided using centrality indices. We will come back to this issue later on but for now it is clear that the two concepts do not only complement each other but furthermore from the perspective of urbanity a combination of the two concepts is required.

Centrality can be considered as the amount of interaction that exceeds the one that would be adequate for a place with certain locational or structural properties of the secondary interaction potential. Centrality thus provides information on hidden pull factors, which can be understood based on complexity theory. It is the mutual reinforcement of different factors as described above that leads to an additional amount of interaction that has seemingly been produced by hidden pull factors. If a place shows a surplus of interaction, we can deduce a high dynamism of that place and even expect urban or urban-like processes. This surplus, which is characterized by centrality, is an indication for a place to be of extraordinary importance within its settlement system, economic system, and cultural system. Centrality is not equal to urbanity, as the term does not cover all aspects of urbanity, but is an essential part of it.

In principle, the two variants of centrality, Christaller centrality and the various types of network centrality, are very similar. However, there are some crucial differences which need to be discussed (Nakoinz, 2019).
First of all, Christaller presents a model or, to be more precise, three models on how to achieve centrality. These models, also known as Christaller’s hexagons, describe an optimized supply network in which, for instance, individual and global distances between interaction partners are kept to a minimum. Although network sciences imply similar structures, these are usually not presented as the main models of a concept. Since Christaller’s models describe networks, they can directly be transferred to networks. In this paper, however, we will not cover Christaller’s models at all but his concept and definition of centrality. Another difference between the two variants of centrality can be found in the principle of optimization. Christaller centers benefit from synergies gained by the concentration of interaction nodes at the centers, while network centers gain their synergies from controlling and concentrating connection edges (Nakoinz, 2019). It follows that Christaller centrality mainly describes properties of a place in a regional network, while network centrality is focused on the structural integration of a place into a multi-scalar network.

Secondly, Christaller provides one consistent explicit definition, while network centrality models cover a multitude of different implicit definitions that are very similar. For Christaller centrality is the relative surplus of importance, which is the basis for our definition of centrality being a relative concentration of interaction. In network sciences, however, the focus lies on connectivity, which is not understood in the narrow sense referring to the nodal degree, but in a more general sense. In his centrality concept Christaller has excluded the influence of primary centrality potentials or primary pull factors and focuses only on secondary potentials or pull factors that are caused by social dynamics. In other words, the absolute centrality was adjusted by effects of primary centrality potentials. In contrast to that, the basic measures from network sciences do not consider primary factors at all and therefore provide absolute measures. For unweighted networks the basic measures provide information on the existing connections and for the weighted graphs, they show how the actual interactions influence the connectivity of a place. Both contain primary centrality potentials as invisible factors. Since we are interested in centrality adjusted by the primary potential, which is the crucial point for urbanity, Christaller’s version is the one we need to consider, when it comes to our present problem.

Thirdly, Christaller offers a centrality measure that is valid only for a specific historical context, while network sciences provide several sound measures. The ideal way would be to simply apply network centrality measures using Christaller’s definition. However, that would turn out to be problematic. In his original study Christaller uses population as a node feature to define the size of a place and he uses another node feature to indicate some kind of absolute centrality. In this connection, centrality is measured as absolute centrality that follows the size of a site (population). In order to use centrality measures as defined by network sciences, we would need to translate the nodal information about the places’ size into structural information from external connections. Since the assumed centrality and the node feature have different units of measurement, this is not done by simply weighting the nodes by size or even by inverting size. Christaller in turn uses the same unit of measurement for both the centrality measure and the size measure. A solution would be to look at networks of different scale. Up to now, we only considered networks in which the nodes represent different places. An intra-site network looking at individuals could provide us with the information we need for the other level. However, since the available data is limited, such method seems rather unrealistic and there is no need for it to be considered. Furthermore, Christaller’s centrality measurement originally focussed on specific network structures that correspond with those structures Christaller proposed for the minimization of transport costs. This star-like network covers the complementary area of a central place.

Both the obvious and the more subtle differences between Christaller centrality and the different network centralities require a literal translation of Christaller’s concept into the context of network sciences. As we have already seen, a direct transfer of concepts and methods would not be feasible and lead to inconsistent results. We therefore assume a system of interaction that is mapped using a network (a simplified model of the interaction system). We distinguish between unweighted graphs, which only represent the structure of the interaction system, and weighted graphs, which also map the flow and actual interaction within the interaction system (for a discussion about the flow of interaction and the term itself we refer to Borgatti, 2005; Taylor et al., 2010). While in the first case the most important information gained is the existence of edges, in the second case the most important information revealed lies in the edge weights. Centrality indices applied to unweighted graphs represent some kind of secondary interaction potential, while centrality indices applied to weighted graphs represent the actual concentration of interaction. In order to obtain the relative concentration of interaction, which corresponds to centrality in the sense of our concept (Nakoinz, 2012, 2019; Knitter and Nakoinz, 2018), we need to subtract the centrality measure of the unweighted graph from the one of the weighted graph.

The previous paragraphs have shown that urbanity, centrality, interaction and connectivity are closely intertwined. This information also includes some hints on why networks play a vital role in the process of urbanity. In this connection, we feel the need to clarify this relationship. A network obviously provides places with additional connectivity, which in turn can lead to an additional amount of interaction. The additional amount of interaction increases the complexity as well as the dynamism and vigor that characterizes urban processes. Although, in principle, networks are no precondition of urbanity, they stimulate urban processes and therefore are tightly linked to urbanity. This assessment of the role of networks is valid for inter-town networks as well as for intra-town networks. Both types of networks lead to increasing interaction, albeit different types of interaction, respectively.

### 3. Method

In order to explore the external conditions of urbanization and urbanity, we need to investigate the real network centrality in...
contrast to potential or structural centrality. According to our theoretical considerations, the real network centrality \((\text{Cent})\) equals the real interaction \((\text{Int})\) between one network node \((i)\) and other nodes, minus the connectivity represented by structural centrality \((\text{Con})\).

\[
\text{Cent}_i = \text{Int}_i - \text{Con}_i
\]

While \(\text{Con}_i\) is a centrality index applied to an unweighted graph, \(\text{Int}_i\) is the corresponding centrality measure for weighted graphs (Borgatti, 2005; Opsahl et al., 2010; Yan et al., 2013). In our case, we use degree, strength, closeness and betweenness, since these measures are very basic, easy to understand and directly lead to an interpretation. Other centrality measures (cf. Bonacich, 1991; Koschützki et al., 2005; Benzi and Klymko, 2013; Agryzkov et al., 2019; Larrañeta et al., 2019; Lv et al., 2019; Skibski et al., 2019) shall be neglected for the moment and can be included at a later stage. While degree only maps the number of connections, strength also includes the intensity of the connections. Closeness is a measure for the reach of a certain node and betweenness maps the interaction control.

This approach is different from most other approaches, since both actual interaction indicators and structural centrality indicators are used. It is also consistent with our theory and there are some additional advantages, which need to be mentioned. Firstly, this approach can be interpreted as the combination of an empirical and a theoretical model, which is a requirement to gain new knowledge (Nakoinz, 2018a, 105). In this connection, the real interaction represents the empirical model, while the centrality index applied to an unweighted graph is a prediction of the structural importance of a place and thus a theoretical model. Secondly, this approach minimizes the edge effects. Spatial network analyses particularly tend to have edge effects, which is due to the fact that peripheral edges cannot have the same amount of structural embeddedness as geographical central nodes. Since both \(\text{Int}_i\) and \(\text{Con}_i\) are affected by this, the edge effect is partly compensated in our equation. This allows for a more extensive interpretation of the network analysis results, as the periphery does not need to be excluded completely.

In order to enter actual values into the equation, we need to define nodes, edges and the flow of our network. The nodes comprise the presumed urban centers and other places. In our case the presumed urban centers are princely seats. Due to a lack of research on ordinary settlements and the fact that we do not have access to a decent data set, the other places are rather difficult to define. However, for our analysis neither the exact location nor the exact inventories are required, unless it was our aim to interpret the minor places themselves. Based on this consideration we can simulate the location of the ordinary settlements. The aim of this simulation is not to capture probable or even real site locations, but to capture the sample point for the aggregation of the fibulae data. However, if we simulate too many places we will end up having too many insignificant fibulae data aggregations comprising only small numbers of fibulae. Therefore, we use a mean distance of 15 km between the different places and a hard-core radius of 10 km, which must not contain any other places. In the princely seat’s hard core the other places are substituted by the princely seats themselves. The fibulae are assigned to the nearest place in the set of simulated places and princely seats, which both serve as network nodes.

Since our aim is to obtain connections at different distances, the conventional solution of using a realistic geographic network that connects natural neighbors only (Delaunay graph) is not considered a decent solution for our problem. Instead, we decided to start with a complete graph at the cost of not being able to produce a readable plot of the network connections. The second step of the analysis consisted in cutting some connections to obtain a semi-complete graph, which is required for some applications. The cut-threshold was arbitrarily set to the value 0.85 of relative proximity, with a maximum value of 1.00. For the analysis we have to keep in mind that the edges of the complete graph have been constructed and are not based on archaeological evidence. This excludes, for example, the calculation of the degree for a complete graph or a Delaunay graph.

The archaeological evidence is taken into consideration when weighting the edges, an approach that is part of an ongoing discussion about the special nature and requirements of archaeological networks (Peeples and Roberts, 2013). In our case, the weighting is based on the theoretical concept of cultural distance. Cultural distance has been shown to be a proxy for interaction (Nakoinz, 2013). The methodology is described in detail in other publications (e.g., Cormier et al., 2017) and, from a network research perspective, it can be understood and compared to similarity networks (cf. Östborn and Gerdng, 2015; Habiba et al., 2018). First of all, the relative amount of each type is sampled for every sample point. This is followed by using this so-called spectrum of types to calculate a distance matrix (Figure 2). For this purpose, the Euclidean distance is used, as the data has been adapted to the Euclidean space. In fact, we use compositional data in which the Euclidean distance is valid (van den Boogaart and Tolosana-Delgado, 2013). The distance matrix resulting from the calculations is then used as an adjacency matrix in order to produce the network. The original distance matrix eventually leads to a complete weighted graph. The cut mentioned above is done by replacing values below the threshold in the distance matrix by zero. Finally, to obtain an unweighted graph, the weights can be replaced by zero. This procedure leads to three different networks:

- Full weighted graph
- Cut weighted graph
- Cut unweighted graph

While the unweighted graph represents the connectivity, the weighted graphs represent the actual interaction. The cut and the complete graph differ in the presence and absence of edges, which is defined by rather low interaction. The complete graph also includes connections that are less intensive, while the cut graph focusses on the more vital edges. Another reason why we cut the graph has been mentioned above. As a matter of fact, some analyses cannot be done using a complete graph. Since the networks only map the input data, which does not lead to any results directly related to our research objective, there is no need...
to plot them. The network graphs are not considered a result of this paper but rather an intermediate state of the data. For our case study only the results of the centrality measures of the network are relevant and shown in the figures.

Since the simple and well-known centrality measures are comprehensible and lead to a straightforward interpretation, we decided to use them in our analysis. Using different types of networks the degree, strength, closeness and betweenness are calculated and presented as mapped symbols, of which the size corresponds to the intensity of centrality. In order to obtain a decent size for the symbols on the map, the values of each category are scaled independently of one another. This means that the symbol sizes of one and the same map, which are located between the different time slices of the same category, can be compared to one another. The symbol sizes outside this spectrum will not be included.

All analyses have been done with R and some additional packages (Csardi and Nepusz, 2006; Knitter, 2017; Nakoinz, 2018b; R Core Team, 2018). The analysis scripts of the case study and the data are provided at https://gitlab.com/oliver.nakoinz/urbanityprocess.

4. CASE STUDY

In our case study we apply the ideas outlined in the previous parts of the paper. We want to show the applicability of our concept and, furthermore, that new insights can be gained by using this approach. Applying this specific approach, we tread
new ground, in which we address the involvement of some kind of middle class in urban processes based on fully reproducible quantitative research.

4.1. Objective
The main objective of this case study is to identify places that show a high degree of network centrality and therefore good conditions for urbanity processes. In this connection, it should be emphasized that we focus on urbanity processes and not urbanization. According to our theoretical background, urbanity is a process in which people permanently have to deal with the challenges and opportunities that come with rising connectivity, growing interaction density and increasing complexity. Since urbanity does not describe a fixed state, the understanding of urbanization as the formation of urbanized places seems to be inaccurate. Instead, we understand urbanization as the existence of specific factors that cause urbanity processes. Looking at urbanization in the context of our case study, we would rather discuss the causes of centrality instead of the location of central places or the degree of centrality of places. In this connection, we look for places which gained a high degree of centrality through network integration, which can be considered a factor of the urbanity process.

4.2. Data
In this case study we apply a very restrictive data set, meaning we only use fibulae. This bears some advantages, such as the comparability with the work of Brun (1988), who also used fibulae to compare princely seats to one another, and a narrow focus of interpretation. The disadvantage is that we are not able to gain a general insight into centrality. The fibulae we used are classified according to Mansfeld (1973). The data stems from the shkr database (Nakoinz, 2013). It was inserted during previous projects and completed during the preparation of this paper. The data used in this study originates from Baden-Württemberg in Germany, with a few additions coming from other countries and states such as Hesse (Germany), Rhineland-Palatinate (Germany), Saarland (Germany), and Alsace (France) (Figure 1).

Distributed to the three phases of the late Hallstatt period, the database includes 272 fibulae from Ha D1, 747 fibulae from Ha D2 and 345 fibulae from Ha D3. Considering the different lengths of the phases, the dominant position of Ha D2 becomes fairly clear. While Ha D1 provides two fibulae per year and Ha D3 six fibulae per year, Ha D2 surpasses that with 37 fibulae per year.

The specific scope of interpretation of the fibulae case study is limited to the middle class and does not include the elites or the poor. We also focus on social ties, rather than religious or economic aspects. Therefore, the fibulae primarily map social developments as part of urbanity processes. This information supplements information from other data. Thanks to proxies of wealth and trade, we already know that some places show some kind of economic connectivity (Nakoinz, 2017b).

4.3. Results
In accordance with the above mentioned categories and the three phases mapped in Figures 3–5, our analysis results in different centrality indices. For the complete weighted graph there are only two maps, as for this type of graph the calculation of the degree would not make any sense. Regarding closeness the values for Ha D1 at Kapf, Ha D2 near Hohenasperg and Ha D3 also near and at Hohenasperg as well as at the Heuneburg are above average (Figure 1). In all three phases, the transport corridors at Rhine and Neckar show medium values. The betweenness of the complete weighted graph highlights the same sites, which are Kapf in Ha D1, the surrounding areas of Hohenasperg in Ha D2 and the Heuneburg in Ha D3. Maximum values are observed for Ha D2.

The degree of the cut weighted graph, which for weighted graphs would actually be indicated as the strength, but for reasons of better comparability is mentioned as degree, does not highlight the princely seats at all, but rather excludes them (Figure 4). The increase in Ha D2 is followed by a decrease in Ha D3. Regarding closeness, the princely seats also seem to be mainly excluded and similar to the closeness of the complete graph, the transport corridors are characterized by values above average. The mean values decrease in Ha D2 and increase in Ha D3. In contrast to the degree, the highest closeness values in Ha D3 can be found at one of the princely seats, namely the Heuneburg. Similar to the closeness, the betweenness shows the highest values in Ha D3, which at the same time are the highest values of all phases at the Heuneburg. Other princely seats do not show values above average. The maximum values for betweenness per phase seem to be continuously increasing.

The last category of centrality measures that we present in this study focuses on the differences between the cut weighted graph and the cut unweighted graph (Figure 5). Regarding the degree the maximum values per phase increase. However, the upper range of the values does not highlight the princely seats. The results show a more general increase of the values, rather than an increase for specific places. The closeness shows an abrupt increase in Ha D3 with a focus on the princely seats Heuneburg and Kapf, with maximum values for the Heuneburg. The betweenness provides outstandingly high values for the Heuneburg and, in second place with much lower values, Kapf. In Ha D2 the values above average do not focus on princely seats, but show a rather general increase continuing the process from Ha D1.

4.4. Interpretation
The interpretation of our results is based on previous knowledge. We are convinced that the princely seats are some kind of central place (Krausse, 2008) and particularly function as network centers (Nakoinz, 2013). We furthermore assume that they are characterized by wealth and trade (summarized Nakoinz, 2017b). This case study is based on the assumption that interaction measures applied to an unweighted graph indicate connectivity and that interaction measures applied to a weighted graph indicate absolute interaction. Furthermore, it is assumed that relative interaction measures indicate centrality in the sense as discussed above. Moreover, we assume that the degree maps the chances for interaction, closeness the possible reaches of interaction and betweenness the interaction control. Taking the actually used data, which is the fibulae, into account, we can
assume that we will gain knowledge on social interaction, rather than economic processes. Considering the fibulae we look at people from some kind of middle class who are able to influence society, but are not part of the power elites. They seem to be people who have certain ambitions and a certain degree of influence, but do not have actual political power. This assumption does not deny the existence of elites. We just assume that the majority of the fibulae neither belong to the elites nor to the poor.

The complete graph results highlight the places with the highest degree of interaction. In contrast, the cut graph shows that the places in which social interaction mainly took place come with very small interaction intensities and become rather isolated when neglecting the existing superficial connections. For closeness and betweenness the results are similar. While the Kapf is highlighted in Ha D1, it is the surrounding area of Hohenasperg that is highlighted in Ha D2 and partly in Ha D3. In the latter phase the Heuneburg is in a dominant position as well.

**FIGURE 3** | Closeness and betweenness of full weighted graphs. Gray points are the simulated sites and the Princely Seats, the red circles indicate the centrality according to the mentioned index, where the size of the circles corresponds with the centrality. The symbols of each row is scaled differently.
FIGURE 4 | Degree, closeness, and betweenness of cut weighted graphs. Gray points are the simulated sites and the Princely Seats, the red circles indicate the centrality according to the mentioned index, where the size of the circles corresponds with the centrality. The symbols of each row is scaled differently.
FIGURE 5 | Relative degree, closeness and betweenness of cut graphs. Gray points are the simulated sites and the Princely Seats, the red circles indicate the centrality according to the mentioned index, where the size of the circles corresponds with the centrality. The symbols of each row is scaled differently.
At first glance, it is particularly irritating that the Heuneburg is missing in Ha D1 and Ha D2. Upon closer inspection, however, this seems to make sense: In times of economic growth the control of the power elites would have left fewer opportunities for a free development of a middle class than in times of consolidation and decline.

This interpretation is further supported by the results of the cut graphs. Regarding betweenness the Heuneburg emerges as the most important point in Ha D3. In contrast, Ha D1 and Ha D2 do not show such extensive betweenness values at all. Looking at the degree or in this case strength and closeness, we notice that the princely seats do not play any particular role. When it comes to the number of vital connections, we can observe a general increase from Ha D1 to Ha D2 and a decrease at a later time. This decrease is obviously compensated by a much higher general reach of interaction indicated by closeness.

Centrality, which according to our definition is the relative interaction, maps the interaction that is added to the already existing interaction determined by the interaction potential. Centrality hence identifies the places that act as actual social stimulators. Compared to the absolute interaction, this surplus of interaction is rather small. However, it highlights the most vital places regarding the social development of the middle class and provides additional complexity.

The continuous increase of the relative degree shows that the decrease of the absolute degree in Ha D3 does not tell the whole story. Most places are even better connected than the interaction potential would suggest and we can assume that the natural development is hampered by external factors which lower the interaction potential. The relative closeness, meaning the relative reach of connections, particularly increases in Ha D3. However, it is mainly the Heuneburg that is affected by that and to a lesser extent the Kapf. The relative betweenness shows a similar pattern, with a general increase in Ha D2 and enormous increases in Ha D3 for the Heuneburg, which showed rather small values in Ha D2.

Regarding urbanity, we thus find good conditions for middle class interactivity in Ha D3 at the Heuneburg and to a lesser extent at the Kapf. However, there seems to be hardly any interactivity at the princely seats further north. This is somehow ironic, as the original function of these places as economic network centers had already moved further north (Nakoinz, 2013) and the southern line of the princely seats had been abandoned to other activities. Perhaps it was this asynchronous social development of the middle class and the economic development of the power elites that stopped the urban process eventually and lead to something which Brun and Chaume (2013) called unfinished urbanization.

Urban places and urban processes attract people, because they offer new opportunities and capabilities (Arponen et al., 2016), even, or particularly, in times of crisis. These opportunities and capabilities, however, come at the cost of instability and unpredictability. The art of successfully managing urban places is the art of finding a balance between the reduction of complexity and the limitation of opportunities through organizational structures, in other words, finding the balance between stagnation and collapse. Networks play a dual role in urban processes. On the one hand, they lead to additional connectivity and thus additional complexity. On the other hand, they provide stability by buffering some of the occurring issues. It becomes clear that networks obviously stimulate urban processes. The mentioned buffer function requires synchronous and interlinked processes. However, as shown in this paper, this requirement was not met in the Early Iron Age of South-West Germany. Different social subsystems developed independently from one another and asynchronously constrained each other. Complexity at that time was high. However, as opposed to what is assumed for urban processes, it was neither possible to reduce complexity nor to benefit from it. Above all, it was not possible to take advantage of the network's buffering capability. The processes required for the adaptation to new conditions could not hold pace with the processes of growing complexity. For the Heuneburg it has already been discussed, if the failure to cope with growing complexity lead to the local collapse in Ha D2 (Nakoinz, 2017b). In our paper we address this problem as well. However, we present a more general picture of it.

Thanks to the detailed analysis of interaction and interaction potentials as well as of connectivity and centrality, we are able to get an insight into the development of the Early Iron Age society and the rather tragic role of urban processes at that time. Theoretical considerations provided us with the necessary tools and terms for a detailed analysis. Without the theoretical considerations we would hardly be able to base our interpretation on different results and indicators. On the contrary, we would perhaps treat completely divergent concepts such as centrality in geography and centrality in social network analysis as equivalent and hence come to wrong conclusions.

Finally, we can answer the question raised in the introduction. Although the middle class was involved in urban processes, for the urbanization as a whole it did not play any important role. And although it potentially could have done so, apparently this middle class had been in the wrong place or lived during the wrong phase, in which synergies with other parts of urban processes or the reinforcement thereof were not possible. At least it seems that this component together with additional capabilities of adaptation and innovation was missing at economic central places and the elites' residences. Sustainable urban processes need a balance between stagnation and crisis and this middle class could have been a providing factor.

5. DISCUSSION

The Late Hallstatt period appears to be a time of challenges and experiments. It is a highly dynamic time in which agglomerations emerge and disappear and it is a time of high complexity, which continuously grows. The climax of complexity is the urban processes, in which complexity is the driving force.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in the gitlab repository https://gitlab.com/oliver.nakoinz/urbanityprocess.
AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This paper has been published thanks to EU OPEN-AIRE funding granted to Francesca Fulminante for the Maria Slodowska Curie Project 628818 Past-People-Net (2014–2016). The preparation of this paper was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation - Projektnummer 290139102-SFB 1266 and Projektnummer 252470382). The finalization of this project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy-EXC-2150-390870439.

ACKNOWLEDGMENTS

We are grateful to Francesca Fulminante for inviting us to this publication. We would like to thank the participants of the course ‘Die ersten Städte nördlich der Alpen? Urbanisierungsprozesse in der Eisenzeit’ (Kiel University 2018, course no 200129) for intense discussions.


Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Nakoinz, Bilger and Matzig. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Patterns of Etruscan Urbanism

Simon Stoddart1*, Alessio Palmisano2†, David Redhouse3, Graeme Barker1, Giorgia di Paola4, Laura Motta5, Tom Rasmussen6, Troy Samuels5 and Robert Witcher7

1 McDonald Institute, University of Cambridge, Cambridge, United Kingdom, 2 Department of Ancient History, LMU Munich, Munich, Germany, 3 Department of Archaeology, University of Cambridge, Cambridge, United Kingdom, 4 Soprintendenza Archeologia, Belle Arti e Paesaggio per le Province di Caserta e Benevento, Caserta, Italy, 5 Kelsey Museum, University of Michigan, Ann Arbor, MI, United States, 6 School of Arts, Languages, and Cultures, University of Manchester, Manchester, United Kingdom, 7 Department of Archaeology, University of Durham, Durham, United Kingdom

This paper examines the patterns of Etruscan urbanism by the innovative use of newly available rural data, employing rank size, and indices of centralization. The detailed case study looks at the development of urbanism of pre-Roman Etruria where both robust and delicate urbanism were present alongside one another. To achieve this end, the paper will draw on the complementary features of two recent articles—Redhouse and Stoddart (2011) and Palmisano et al. (2018)—to provide a synthesis that both examines the large places and the supporting rural settlement. The territorial boundaries of the major urban places were predicted by the XTENT model in the first article. The cumulative numbers of rural settlement (and other proxies of population) over time were examined in the second article. This paper will look at the regional variation in landscape organization within the predicted territorial boundaries of the major robust centres and the more delicate transitory centres, as well as the buffer zones in between. At least three phases of boundary development can be examined, equivalent broadly to the Iron Age, Orientalizing/Archaic and Post Archaic periods, seeking to match these with the correspondingly dated rural settlement. The results will be critically examined in terms of broader knowledge of the economic and political development from current fieldwork in Etruria. The ethnographic analysis of Kopytoff (1989) will also be applied to assess the application of the internal African frontier to the central Italian context. In this way, the quantitative will be matched with the qualitative to provide a deeper understanding of urban development in an under-assessed example within the Mediterranean world.

Keywords: urban—rural, urban development, Etruria, Mediterranean, city

INTRODUCTION

Background
The Etruscan settlement pattern analyzed here belonged to rich communities living in an area generally defined as Tyrrenian central Italy during the first millennium BC. These were the communities that competed with the Latins to the south of the Tiber, and were later absorbed into the Roman empire (Figure 1A). That means that they lived principally between the river Arno to the north at 44° North, the river Tiber to the south at 40° North and toward the south east, and the Mediterranean to the west.
FIGURE 1 | (A) Map illustrating the location of Etruria and the places and regions mentioned in the text. (B) The region of Etruria, showing the survey zones and the recovered rural settlement. The numbered surveys are indicated as follows: (1) Nepi Modona (1953); (2) Perazzi and Poggiesi (2011); (3) Chellini (2012); (4) Valenti (1999); (5) Valenti (1995); (6) unpublished, courtesy of Di Paola; (7) Cucini (1985); (8) Curri (1978); (9) Campana (2001); (10) Cenni (2010); (11) Felici (2012); (12) Campana (2013); (13) Felci (2004); (14) Paolucci and Francovich (2007); (15) Botarelli (2004); (16) Cambi (1996); (17) Carandini et al. (2002); (18) Quilici Gigli (1970); (19) unpublished, courtesy of Barker and Rasmussen, 1998; (20) Corsi (2000); (21) Nardi (1980); (22) Quilici Gigli (1978); (23) Hemphill (2000); (24) Andreussi (1977); (25) Morselli (1980); (26) Rajala (2013); (27) Gianfrotta (1972); (28) Enei (2001); (29) Tartara (1999); (30) Verga (2006); (31) Muzzioli (1980); (32) Patterson et al. (2020).
At the beginning of the Early Iron Age (1,020–900 BC), small dispersed sites (ca. 2–3 ha, up to 15 ha) often on turf outcrops were abandoned and new large proto-urban centres on larger plateaux (ca. 50–185 ha) were occupied (Guidi, 1989, 2010; Pacciarelli, 2000; Poroni, 2000; Bonghi Jovino, 2005; Milletti et al., 2010; Fulminante, 2014, p. 44–47; Alessandri, 2015, 2016; Stoddart, 2016). In the Late Iron Age/Orientalizing (750/725–580 BC) and Archaic periods (580–480 BC), the settlements reached full urbanization and the political landscape was divided into several competing city-states distributed at an average radial distance of 15–25 km, accompanied by increased intensity of rural settlement.

The position of Etruria in the central Mediterranean gave these communities privileged access to the efficiencies of maritime transport and trade, once technological advances in shipping had reduced the risks of sea travel to a level that granted these communities considerable advantage. The benefit of communication was enhanced by access to at least four significant rivers penetrating into the intermontane valleys of the foothills of the Apennines: the Tiber, the Arno, the Ombrone, and the Albegna. These rivers assisted the extraction of resources from the mountainous uplands of the Apennines, providing a complementarity between the fertile volcanic agricultural soils and metal resources of the relative lowlands and the pastoral zones of the uplands. In this way, geopolitics underlay the landscape configuration of the Etruscans and this paper explores and extends that logic.

The distinctive urban culture of the Etruscans has long been known as an elite strategy, understood through material culture, particularly the study of pottery, and larger scale visual culture (Stoddart, 2020a). This is the rich analysis undertaken by a series of distinguished authors (e.g., Ampolo, 1980; Bettelli, 1997; Bartoloni, 2003; Smith, 2014; and many others), but that is not our approach here. The underlying settlement structure that supported this urban culture has been less explored, in spite of the considerable potential availability of settlement data from the study of cities and cemeteries since the nineteenth century, and from regional landscape survey concentrated in the second half of the twentieth century.

More recently, the synthesis of the promising urban landscape data has been made possible by a series of converging factors: a strong survey tradition, the provision of raw data from these surveys and the modeling of city territories. Early steps before the fully digital age (Stoddart, 1987; Rendeli, 1993; Pacciarelli, 2000; Cifani, 2003) have provided a foundation for the full assembly of published survey data (Palmisano et al., 2017, 2018), now for the first time combined with four other prominent systematically collected survey data: Cecina valley (Samuels and Terrenato, unpublished data), Populonia (Di Paola, 2018), Tuscania (Barker and Rasmussen, 1998), and the Tiber Valley Project [British School at Rome, Patterson et al. (2020)] yet to be fully published. Whereas, syntheses have been assembled and interpreted for the subsequent Roman period (e.g., Patterson, 1987; Launaro, 2011; Sewell and Witcher, 2015), the same analysis has never been achieved for the formative Etruscan phase and this is the focus of the current urban analysis, one that deserves comparison with many other classic case studies of state formation in the ancient world.

The study of the apparently dominant centres of power (Redhouse and Stoddart, 2011; Stoddart, 2020b) and their implications for the differential organization of the countryside has already been achieved. For this preceding study, we have what amounts to a nearly complete sample, since no sampling strategy can easily ignore the major urban foci surrounded by their cemeteries (Dennis, 1848). However, even here some caution should be observed about the completeness of archaeological samples, since some very large urban centres have been discovered in relatively recent times: e.g., Doganella (Perkins and Walker, 1990), Marsiliana d’Albegna (Zifferero, 2010), Gonfienti (Poggesi et al., 2005). Inevitably, some characteristics of this sample are open to debate, including even basic features such as size, internal density and thus population, but we do have a fairly complete data set of settlements larger than 10 ha.

What has never been achieved before is to combine these powerful places with the rural settlement from their territories. These amount to 1894 sites dating from the ninth to the fifth century BC, uncovered by systematic field survey. The combination of these two data sets provides an innovative comparison of the strategies of the individual urban centres. This, in turn, provides the living backdrop to, and infrastructure for, the different urban cultures which have been known for a much longer period, as first recognized by Banti (1960). In spite of the relatively large data set, as we discuss further below, we do not have anything that can be characterized as a complete sample, and the source criticism of any such archaeological data set is an important prererequisite before drawing any sound conclusions.

The careful characterization of these data nevertheless allows new interpretations of the multiple strategies employed by the citizens of urban Etruria. The urban centres were in broad equilibrium until the sack of Veio and the arrival of Rome, but they varied considerably in size of centre, size of territory, and density of occupation of the landscape. One important outcome is to be able to outline the dynamism of the urban landscape. The central points of the landscape were generally long-lasting (“robust”; cf. Stoddart, 2016), but as these central points strengthened they prevailed over frontier areas that were more delicate in their organization. This robust urbanism is based on urban centres which endured for at least 400 or 500 years. This is an urbanism with well-defined practices, often institutions, which passed on authority and power from one generation to the next beyond the memory and charisma of the individual, apparently without repeated crisis. In contrast, we can observe what Stoddart has termed “delicate” urbanism (Stoddart, 2017), involving short-lived urban centres which lasted little more than a few generations and perhaps lacked well-rehearsed practices for passing on power and authority or, if attempting to develop comparable practices, were caught in a geopolitical trap, pressured by their nearest large neighbours. Such delicate urban centres may be relatively substantial, but the nucleated populations dispersed after relatively short periods of few hundred years. In Etruria, both forms—robust and delicate—were, for a time, contiguous, and direct comparison between them can be made within a broadly similar cultural landscape.

In this paper, our analysis has been achieved by combining several techniques in combination as explained in more detail below. First, hypothesized territories have been defined for
the major urban centres using the mathematical modeling of XTENT (Renfrew and Level, 1979) sensitized to the physiography of the landscape (Redhouse and Stoddart, 2011; Stoddart, 2020b). Secondly, the rural settlement has been allocated to these territories and the degree of hierarchization, centralization, and density of settlement assessed quantitatively. Three broad chronological phases of boundary development will be examined, equivalent broadly to the Iron Age/Orientalizing (ca. 1,020/1,000–580 BC), Archaic (ca. 580–480 BC), and Post-Archaic (ca. 480–350 BC) periods, seeking to match these with the correspondingly dated rural settlement. These time frames have been chosen because they reflect the level of accuracy of datable surface material—mainly pottery collected from the ploughsoil by different survey projects—although they are of unequal length. The results show the multiple strategies employed by the agents of urbanism in Etruria.

The Archaeological Data and Case Study

The high intensity of field archaeological investigation (e.g., excavations, surveys) in a relatively dry arable landscape makes central Italy an excellent case study for assessing demographic trends (Barker, 1988; Potter and Stoddart, 2001; Stoddart, 2007), and the first millennium BC which forms the focus of this study has a consistency of recovery only exceeded by the Roman period.

For many years, these projects have remained separate datasets. This situation changed when Palmisano et al. (2018) collected a comprehensive dataset of archaeological sites in central Italy. For this, they conducted a comprehensive review and harmonization of settlement data from archaeological survey reports covering an overall extent of ca. 10,000 sq. km and a chronological scope spanning from the Late Mesolithic to the fall of the Roman Empire (ca. 8000 BC–500 AD).

For the current exercise in the restricted period (first millennium BC) and region (Tyrrenian central Italy), this data set has been augmented by the inclusion of new data from published sources and four important unpublished data sets kindly made available for this publication leading to a final total of 549 Iron Age/Orientalizing, 1,248 Archaic sites, and 914 Post-Archaic sites. Settlement data were recorded as georeferenced points (unprojected WGS84). In the present paper, the term "site" refers only to those places identified as dwelling places by excluding cemeteries, temporary activity areas (e.g., campsites) and industrial zones (e.g., mines). One major new set, that of Veio (Tiber Valley Project), lacks the systematic measurement of size and therefore the new data cannot be included in the full set of analyses. We remain dependent on a smaller set of size data from this territory (Rajala, 2013). Further work could also be undertaken on local publications particularly those of local groups, the grey literature of the Superintendencies and by including the well researched, empty, parts of landscape (e.g., Ceccarelli and Stoddart, in press), but this data set nevertheless comprises the most complete data set yet assembled of Etruscan rural data. The archaeological settlement data used in this paper are presented spatially in Figure 1B.

These data sets must necessarily be approached with caution since they conceal multiple methodological strategies for the recovery of archaeological material with varying levels of interest, knowledge and specialization and differing local site formation processes. These issues have been extensively studied by fieldworkers (di Gennaro and Stoddart, 1982; Stoddart and Whitehead, 1991; Terrenato, 1996; Binliff and Sbonias, 1999; Francovich et al., 2000; Mattingly, 2000), but will not be deeply analyzed here. We nevertheless maintain that the large size of the data set does enable the detection of trends even at the more localized level of comparing different territories.

The most important differences to be noted are the degree to which individual city territories have been comprehensively covered by survey. It is rare to find a strategy of 100% coverage for all sorts of practical reasons. Veio, Nepi, and Murlo stand out for their more comprehensive coverage, which is substantially complete. The coverage around Cerveteri, Tuscania, Marsiliana d’Albegna, Doganella, and Chiusi is extensive, but incomplete, often structured by the choice of survey area and project sampling and collection strategies, matters explicitly set out in the survey reports. Other urban centres such as Acquarossa, Populonia, Vetulonia, and Fiesole have less detailed coverage, often built up from many sources of information.

METHODS FOR DEFINING TERRITORIES, SETTLEMENT HIERARCHIES, AND REGIONAL CENTRALIZATION

XTENT Model

As outlined in the 2011 paper (Redhouse and Stoddart, 2011), the XTENT model was devised by Renfrew and Level (1979) to overcome a number of simplifications faced by the Thiessen (1911) polygon or Voronoi polygon analysis used in the original Early State Module (ESM) analysis of Renfrew (1975). The main simplification is that the original ESM analysis divided up territory equally between centres regardless of size or of any physical barriers in the landscape. Instead, XTENT is based on the simple assumption that territorial extents are related both to the size and the distance between urban centres. In particular, the influence exerted by a centre on a specific location of the landscape can be modeled according to the relative size of that centre and its distance away from that given location. Therefore, this technique permits the prediction of buffer zones of unallocated political space that can be detected cross-culturally in developing political landscapes (cf. Marcus and Feinman, 1998). In the present analysis, the calculation of the territorial extent is measured against the “friction” of the physiography of the landscape by introducing a Digital Elevation Model (DEM) into the equation (cf. Duve and Kroefges, 2008; Bevan, 2010). In this way, the formula of XTENT is contextualized to the specific physical environment.

In the particular case of Etruria (Figure 2), the methodology executed was as follows. Within the boundaries of the sea and the River Tiber (an important cultural boundary), political boundaries were calculated in all directions from each major primate centre using the following mathematical reasoning:

\[ I = f(C) - k \cdot d(I \geq 0) \]
Where \( I \) is a measure of influence at a given location, \( C \) is a measure of size of the centre, \( d \) is a measure of distance from the centre, and \( k \) is a constant. A centre \( C_1 \) will dominate a centre \( C_2 \) if \( I_1 > I_2 \) at \( C_2 \), that is if:

\[
 f(C_1) - f(C_2) > k \cdot d_{1,2}
\]

The constant \( k \), representing the fall-off of influence was, following Renfrew and Level (1979), investigated empirically, using their suggestions and previous experience with the technique. In the original unpublished analyses by Harrison and Stoddart, the Distance \( d \) was measured as a simple linear distance. In the revised analysis, the distance was transformed to register travel time by taking into account varieties of terrain.

The constant \( k \), representing the fall-off of influence was, following Renfrew and Level (1979), investigated empirically, using their suggestions and previous experience with the technique. In the original unpublished analyses by Harrison and Stoddart, the Distance \( d \) was measured as a simple linear distance. In the revised analysis, the distance was transformed to register travel time by taking into account varieties of terrain.

The first equation can in that case be written as:

\[
 I = f(C) - k \cdot d \cdot w(I \geq 0)
\]

Where \( w \) represents a transformation to take account of varieties of terrain. Implementing this model requires the following:

1. A tool that will determine the cost of traveling from a site to any point within the area of interest, in other words a routine for calculating \( d \times w \)
2. A tool that will calculate the value of \( I \), based on the value of \( d \times w \)
3. A tool that will determine based upon the above, for a group of sites, which has the greatest \( I \) at a given point.
The ArcInfo GRID function pathdistance() (Environmental Systems Research Institute, 2001) calculates a least-accumulated cost model accounting for surveyor's distance and horizontal and vertical cost factors. This provides us with, for a given location within the area of interest, the value of \( d \times w \) with respect to a particular site. Simple map algebra within GRID permits the calculation of \( I \) at all locations within the area of interest, for a given site. The ArcInfo GRID function upos() (Environmental Systems Research Institute, 2001) generates from a set of input grids an output indicating which grid has the highest value at a given location. The procedure requires as inputs a list of archaeological sites with co-ordinates, site names, site sizes (C), and a DEM. A least-accumulated cost model, in the form of a grid extending to the limits of the area of interest, is generated for each archaeological site. The influence \( I \) for each site is calculated using the least-accumulated cost models. A grid covering the area of interest with all cells \( = 0 \) is also generated. The grids containing \( I \) for each site, and the grid consisting entirely of zeroes are used as inputs to the upos() function.

The resulting output records the site with the greatest value of \( I \) at each location, or zero if all of the sites have a negative \( I \) at a particular location. The least-accumulated cost model does not presently incorporate any consideration of rivers, lakes, and the coast. The least-accumulated cost model only considers the cost of traveling away from an archaeological site. It should also be recognized that the resolution of the DEM was only 80 m. Further development of the model could take these factors into account. Nevertheless, we think that this resolution works reasonably well at the regional scale of analysis.

In the present exercise, the latitude and longitude were established using the Getty Thesaurus of Geographic Names Online or calculated at an appropriate level of precision from a map. As Etruscan specialists will know, there is much controversy considering the values of site sizes for Etruscan and Umbrian sites founded in the period 900–600 BC, since many of the estimates are based merely on topographic location. Where possible, the starting nucleation point in the Iron Age was taken, drawing on Pacciarelli (2000) for up-to-date consideration of this issue. Comparison was also made with the two main rank size studies of Etruria (Judson and Hemphill, 1981; Guidi, 1985) and, where information was still not forthcoming, the estimates provided by Mansueli (1985). These estimates still left considerable gaps, and estimates have been made for Murlo, Pisa, Gubbio, Assisi, Città di Castello, Spoleto, and Todi based on personal experience. Given the fieldwork strategy at Murlo and the heavily wooded terrain it is difficult to be more precise. The estimate for Pisa recognizes the considerable fieldwork undertaken there in recent years, although it is very difficult to be precise about site size. The estimate for Chiusi has increased in recent years and the figure employed here is a compromise figure to reflect the greater understanding of its earlier history, including its polyfocality, than was understood before. The Umbrian site territories are of repeatedly similar size in lake basin catchments so the estimates give a reasonable illustration of the settlement process. The nature of the primate organization of Etruscan settlement permits the analysis of relatively small number of sites since these were generally dominant in their landscape, and the results would not have been affected by the introduction of the rural sites now available. However, there are some sites which could be included in further work, including Gualdo Tadino and Colfiorito on the Apennine margins, and, more importantly, Amelia and Terni in southern Umbria. One great advantage of XTENT is that the mapping can be repeated using different values to explore the consequences, and clearly the results of these changed values should be implemented in future work. These sizes were assimilated with those provided by Palmisano et al. (2017, 2018) and combined with the four unpublished data sets, and the data deployed are included as supplementary online open access information (where the contributing authors are ready to release them).

In any spatial analysis, boundary issues are a key consideration. The area of study was defined by the Tyrrhenian sea to the west, and the Apennines to the north and the east. To the south, the Tiber was taken as the key cultural boundary, and thus the effect of Latin and Sabine centres was not considered. Equally as already mentioned above, southern Umbria was only partly included in the analysis and the centres of Terni, Amelia, Narni, and Otricoli have not been included, but their presence has been predicted by the XTENT results as a political vacuum, filled by Veio in their absence. After these decisions, the area of interest for the purpose of the XTENT analysis was defined by buffering the sites at \( 2 \times \) the mean nearest neighbour distance and then clipping this polygon with the Italian coastline and the course of the Tiber. The DEM was created by obtaining elevation data from the Shuttle Radar Topography Mission (Rosen et al., 2000) transforming it to a Transverse Mercator projection re-projecting it in point form to IT_ED50/UTM, and then generating an elevation model using topogrid (Environmental Systems Research Institute, 2001). The energy-cost model was based upon the measurements of the cost of traversing slopes by Minetti et al. (2002).

**Rank-Size Analysis**

The classic geographical approach to rank size is to plot the rank of sites against their size on logarithmic axes. Modern urban geographers noted that in well-developed urban systems this produced a straight (or lognormal) line, following the so-called rank size or lognormal rule where the second ranked site was half the size of the largest, the third ranked sites was one third the size of the largest, and so on (cf. Auerbach, 1913; Zipf, 1949). This simple observation has gradually attracted both interpretations and quantification. Interpretation has emphasized the tendency of mature urban systems toward the lognormal (Savage, 1997; Cristelli et al., 2012; Fulminante, 2014; Jiang et al., 2015). Rank-size graphs are plotted on a logarithmic scale and the Zipf's Law appears as a straight line from the upper left to the lower right corner of the plot (Figure 3AA). Settlement size distributions of archaeological sites rarely conform to the Zipf's Law and they can be shallower (convex distribution, Figure 3AB) or steeper (primate distribution, Figure 3AC). Heavily centralized urban systems, characterized by one or very few large centres and many smaller settlements, are considered primate (Johnson, 1977; Paynter, 1982; Ades and Glaeser, 1995; Falconer and Savage,
Moreover, one very real problem of rank size analysis is that its results are profoundly affected by the boundaries of the system under analysis. This is a problem in archaeology because it is impossible to define the exact boundaries of a past polity or settlement system. In addition, the spatial scale of analysis affects the results in different settlement size distributions (cf. Drennan and Peterson, 2004, p. 535–539; Palmisano, 2017). Hence, we deploy XTENT as one solution to this problem, since XTENT defines the major, most probably politically independent, nucleations in any given landscape, and provides territories for them by following clearly specified principles.

One of the most useful aspects of the rank size analysis is as a measure of centralization, and a series of quantitative measures have been developed by Drennan and Peterson (2004) to establish indices of centralization and the degree of significance, an update on the earlier statistical Rank Size Index (RSI) (Johnson, 1981). They propose an A-coefficient, which calculates the proportion of the area between the Zipf’s Law line and the observed rank-size curves (see also Crema, 2013, 2014; Palmisano, 2017 for the application of this method; Figure 3B). Hence, the area above and below the Zipf’s Law will have, respectively, positive (A1) and negative (A2) values (Figure 3B). The maximum value for A1 is by definition 1, while A2 could exceed −1 for strongly primate systems where one or more observed settlements are smaller than the expected smallest settlement predicted by the Zipf’s Law. As a result, convex curves will show positive A-coefficient values, while primate patterns negative values. In addition, a bootstrap method has been enabled to test the statistical significance of the A values (cf. Drennan and Peterson, 2004, p. 539–543). This technique calculates the 95% confidence interval of A values by resampling with replacement the observed settlement sizes with 1,000 samples randomly selected. By way of illustration, the rank-size of a putative XTENT-defined territory with 20 sites would be repeated with 1,000 random samples of 20 sites that can be compared with the observed original dataset. In this way, alternative patterns can be tested against the observed patterns. In each graph, the simulated samples (grey lines) are plotted against the observed patterns (dark line), such that a narrower envelope emerges for more certain outcomes and a wider envelope for less certain outcomes.

Measuring Centralization
A further analysis for measuring regional centralization consists of calculating the proportion of the total population within each concentric ring (or “donut”) radiating from a given urban centre. This exercise allows the calculation of the B-coefficient (see Drennan and Peterson, 2008), where the B value ranges between 0 and 1 (0 = no centralization at all; 1 = maximum centralization). The B-coefficient is calculated as follows. First, the territory of a given polity is divided into 10 concentric rings moving away from a given urban centre at a fixed distance of 1 km between rings. In the strongest possible centralized scenario, the innermost ring would contain the 100% of population (or the total estimated settlement’s size) and the sum of the cumulative proportions would be 100 × 10 (n. of rings) = 1,000. In a non-centralized settlement system, the population would be distributed evenly and each ring would contain the 10% of the
polity’s population (or total estimated size) and the sum of the cumulative proportion would be 550 (= 10 + 20 + 30 + 40 and so on). The difference between the sum of cumulative proportions with maximum centralization (1,000) and no centralization at all (550) by using 10 concentric rings is 450. Therefore, the B-coefficient is calculated by subtracting 550 to the sum of the observed cumulative proportion and dividing the remainder by 450.

We used the R free software environment (version 3.6.1) for performing the rank-size analysis and measuring regional centralization (see Appendix A).

RESULTS

Mapping the ESM Model

One of the predictions of the ESM model is that Voronoi [or Thiessen (1911)] polygons might represent the territories of the individual centres, and that an area of 1,500 sq km would be a likely surface area of individual territories (or polities in later literature). As a first step toward the analysis, the areas of the territory of individual centres were calculated on this basis in the 2011 article. The size of the territories relates to the level of packing (space for territory) in the landscape. As a consequence, in South Etruria, very few centres reach the 1,500 sq km threshold. Of the two that do, Veio is very understandable, but the power of Civita Castellana (Falerii Veteres) is greatly increased compared with expectations (although rivals such as Nepi and Narce were not included in this analysis). In North Etruria, the more spacious conditions permit seven out of 11 centres to achieve the predicted territory size. However, although Volterra’s status is quite understandable, the role of Fiesole, Murlo, and Pisa was greatly increased compared with expectations. Interestingly it is the Umbrian fringe that conforms most consistently to the predictions, because it is here that equal spacing is most consistently followed in a sequence of lake basins. The match would probably be even better if Gualdo Tadino and Terni were brought into consideration (with a corresponding decrease in the area of Gubbio and Spoleto).

Presentation of the XTENT Results

For the comparative and heuristic purposes, we outline here the essence of the original 2011 article (Redhouse and Stoddart, 2011). Renfrew and Level argue that variations in the constant value k permit the mimicking of the developing political landscape and this principle was applied to Etruria. In this analysis nine values of k (0.1, 0.08, 0.07, 0.055, 0.03, 0.02, 0.018, 0.016, and 0.014) were applied while holding the f(C) constant at 0.5. For simplicity, three values of k (0.055, 0.03, and 0.016) are again presented here, although the full plots are represented elsewhere (Stoddart, 2020b). Higher values present very small territories. Lower values present the collapse of the political structure of the landscape, ominously suggesting the encroachment of power from the south by Rome, here represented by Veio, since Rome is not considered in the present exercise, and, if included, its size would have presented an even greater threat under the conditions of the XTENT model.

The complete sequence of three “phases” (Figures 2A–C) shows a number of interesting developments: the emergence of corridors of political vacuum, often anchored on river valleys, lakes, and prominent mountains; the survival of key intermediate-sized centres in the interstices between the major centres; a more rapidly maturing political landscape in the south of Etruria compared with the more widely spaced north; and a contrast between the developing disparities of territory size in the south and the regularities of territory size in the inner parts of Etruria and Umbria. For prehistorians, it is also useful to note that the independence of these large primate centres can be predicted from the use of XTENT without any resort to literary sources (cf. Spivey and Stoddart, 1990). For the purposes of the analysis in the present paper, the Iron Age/Orientalizing is hypothesized to match K = 0.055, the Archaic K = 0.03, and the Post-Archaic K = 0.016 (see Figure 2).

Under this analysis, the individual territories of Etruria and Umbria present some strikingly different trajectories. In South Etruria, three megacentres (Veio, Orvieto, and Vulci) present expansionist trajectories, although only Veio appears unstoppable in its development. History, of course, checked this occurrence by action from south of the Tiber, in the form of Rome. The territorial development of Tarquinia, and even to a greater extent of Cerveteri, was checked by enclosing polities. The check was such that Cerveteri could not achieve the 1,500 sq km threshold unlike all the other large centres. Finally, three centres, Acquarossa, Civita Castellana, and Bisenzio were eliminated by their larger rivals during the Archaic (ca. 680–480 BC) and Post-Archaic (ca. 480–350 BC) periods. Of these predictions only that of Civita Castellana is controversial and is discussed more below. The contrast between the life histories of different centres and their accompanying territories is made explicit by the different developments of Acquarossa and Veio.

In North Etruria, long term development is dominated by the expansion of Volterra, unrivaled in its control of the hinterland of North Etruria and able to penetrate to the sea through the Arno valley during the Archaic and Post-Archaic periods (Figures 2B, C). Five other centres (Chiusi, Vetulonia, Populonia, Arezzo, and Fiesole) conform very well to the ESM predictions, and indeed enter some form of equilibrium just above the 1,500 sq km level. Four other centres (Murlo, Roselle, Cortona, and Pisa) were eliminated or squeezed from the political landscape. Of these results that of Roselle is the most controversial and discussed more below. The contrast in trajectories is made clearest by comparing Chiusi’s expansion with Murlo’s decline during the Archaic and Post-Archaic periods. In eastern Etruria and Umbria, there is much more of an equilibrium. Five out of the six centres settle at a level in the region below the 1,500 sq km prediction of the ESM model, although there is some variation between Assisi, hemmed in by rival polities, and Gubbio with less marked constraints. The only centre whose demise is predicted is that of Todi under pressure from Orvieto from the east. An interesting contrast is visible in the development of Gubbio and Perugia. Perugia’s territorial development is forced into a plateau, restrained by the packed nature of her political environment. Gubbio, still one of the largest modern local government districts today, had greater freedom to expand, only limited by the
Apennines to the north and east (although Gualdo Tadino to the east was not considered in this analysis and may have provided check in this direction).

A more detailed analysis of the plot produced by the $k = 0.03$ (Figure 2B) value demonstrates the potential of the integration of XTENT as a heuristic technique for confronting archaeological and historical information. It is the errors as much as the predictions that are insightful. A brief analysis of the landscape from south to north reveals the following observations which build on an earlier analysis (Stoddart, 1990; Redhouse and Stoddart, 2011). The technique suggests that Veio is cut off from the sea by the territory of Cerveteri and that the Faliscan territory is taken over by Veio to the north during the Archaic Period (ca. 580–480 BC). These are both widely debated political issues. Some authors emphasize the overwhelming power of Veio (di Gennaro and Schiappelli, 2004; Cifani, 2005) whereas others prefer to emphasize the independent identity of the Faliscan territory (Colonna, 1990). This is a debate between demographic and spatial logic, and cultural and historical tradition. Further north, the patterns make good, and uncontroversial, geographical, and historical sense. The territory of Tarquinia neatly fits the catchment of the Marta river and is restricted by the medium-sized centres of Bisenzio and Acquarossa in the hinterland. This part of the landscape has encroached on the Albegna valley to the north west, and is restricted by the medium-sized centre of Bisenzio and the high ground of Monte Amiata. This territory adjoins the border region of the Albegna valley which was a zone of instability in the early development of the Etruscan landscape, a point noted by the presence of unallocated political corridor when higher k values of XTENT are applied. It is in this corridor that, first a series of small unstable settlements—e.g., Marsiliana d'Albegna (Zifferero, 2010)—developed, and then the massive entrepot of Doganella (Perkins and Walker, 1990) precisely on the XTENT boundary.

Further north, the territory of Vetulonia has overaken the territory of Roselle. In historical reality, the centre of Vetulonia went into decline and Roselle took over the more prominent role in the local area. This is the one instance where the results of XTENT clearly contravene the patterns of historical development, because the general patterns of spacing of primate centres are also contravened in this case. Local political conditions led to the changed concentration of power in these two centres. One contributing reason may be the importance of lagoonal areas and of the local river (Ombrone) in the development of Roselle and its communications with the interior. These are factors not considered in the present analysis. Another interesting prediction lies in the penetration of Populonia's power up the coast of Etruria into the Cecina valley (exploiting the low relief up the coast) and threatening Volterra's access to the sea. This is clearly another buffer area of unstable political centres, particularly during the Orientalizing period in an area which has been subject to recent field research (Terrenato, 1992; Regoli and Terrenato, 2000).

In the inland area of Etruria, three political territories and one upland area from south to north, Acquarossa, Bisenzio, Monte Amiata, and Murlo form a buffer zone between the coastal states and the inland states. Interestingly this buffer strip converges and overlaps with the line of volcanic lakes Bracciano, Vico, Bolsena that straddle the political boundaries to the south and, at a smaller scale, are also associated with small boundary centres such as Grotte di Castro. Behind this screen of political centres threatened by larger neighbours, there is the final large scale territory of Orvieto. Only this centre, straddling the river valley to the north, has the same scale of territory as some of the territories of its coastal rivals. This centre is crucial in providing the corresponding political pressure on Bisenzio and Acquarossa which went into decline and were replaced by the formal ritualization of political boundaries (Riva and Stoddart, 1996; Zifferero, 2002) once these centres had been absorbed into the larger territories.

Murlo to the north is a more controversial centre (Phillips, 1970; Cristofani, 1975; Torelli, 1983; Stoddart, 1995), in part because of the particular methodology of its exploration, in part because of a smaller scale of political operation. Recent work has shown that the centre, extensively explored as a source of material culture (Phillips, 1970, 1993; Phillips and Talocchini, 1980; Tuck et al., 2006; Shipley, 2017), did not operate in isolation (Campana, 2001) and was part of a small scale local network of sites (incorporated in the analysis below). This centre was also on a sensitive political boundary that stretched north to Castelnuovo di Berardenga (Mangani, 1985) and south to Poggio Civitella (Donati and Cecarelli, 2002). The status of this boundary changed from a string of independent political entities (in the Orientalizing and Archaic) to a series of fortified sites between major political states (in the Hellenistic period); this phenomenon is best indicated by the changed role of Poggio Civitella to the south from small settlement to fortress, but also by the foundation of other fortified sites in the north Chianti region (Becker, 2002, 2008) that replaced sites such as Castelnuovo di Berardenga.

A series of sites—Chiusi, Cortona, Arezzo, and Fiesole—occupied the communication route up the Chiana river and extended along the Arno river reaching, after a further phase of political expansion, to Fiesole which in turn guarded access through the Apennines to the Po valley by means of the Mugello. The newly discovered Etruscan site at Gonfienti (Poggesi et al., 2005) appears to be yet another short-lived “boundary” site in succession to earlier centres at Artimino and Quinto Fiorentino. Although there is some recent discussion over the size of Chiusi (Cappuccini, 2010), which shows signs of expansion under the current reconstruction, all these centres were relatively small compared with the centres to the south and west. A combination of dense packing in the available space and size has led to a distinctively different arrangement of territories. To the east, a prominent landscape feature, the lake of Trasimene, again acted as a frontier. In this case, the lake was bisected by the boundaries of three states which underwent a comparable transition towards ritualization in the later phases (Paolucci, 2002). Perugia, the frontier Etruscan city (Ceccarelli and Stoddart, in press), supported by its greater demographic weight, and facilitated by the morphology of the Valle Umbra, projected east with a slightly larger territory, surrounded by smaller "Umbrian" neighbours. To the north and east, the Apennines provided a distinct physical boundary nicely emphasized by XTENT. In the south of Umbria,
FIGURE 4 | Histograms of site size frequency for Etruria by period. (A) All sites in the Iron Age and Orientalizing. (B) Iron Age and Orientalizing sites measuring 0–50 hectares. (C) Iron Age and Orientalizing sites measuring 50–200 hectares. (D) All sites in the Archaic Period. (E) Archaic sites measuring 0–50 hectares. (F) Archaic sites measuring 50–200 hectares. (G) All Post-Archaic sites. (H) Post-Archaic sites measuring 0–50 hectares. (I) Post-Archaic sites measuring 50–200 hectares.
FIGURE 5 | Histograms of settlement size for each city territory for the Archaic period. (A) Volterra, (B) Fiesole, (C) Murlo, (D) Populonia, (E) Vetulonia, (F) Chiusi, (G) Vulci, (H) Acquarossa, (I) Orvieto, (J) Tarquinia, (K) Cerveteri, (L) Veio.
there appears to be a political vacuum which would be occupied by Terni were it to be included in the analysis. The inclusion of Terni and Amelia, as well as other smaller centres, would also block the advance of Veio into this area predicted by XTENT without their inclusion.

**Settlement Hierarchy From Simple Site Size Histograms**

The next step has been to provide simple presentational statistics of the site size frequency for the whole of Etruria combining the primate centres (from XTENT) and rural farmsteads (from the systematic surveys) (Figure 4). The results show the contrast between the profusion of small rural settlement and the small number of larger centres, already suggesting substantial primacy and regional centralization in the configuration of urbanism. The subsequent step is to show the same data (where size data have been collected) for each of the territories (Figure 5). These generally confirm the dominance of the primate centres during the Archaic period, which then evolved into full-scale urbanized societies. This analysis also shows that some of the data sets (Fiesole, Vetulonia, Orvieto, Tarquinia, and Veio) have potentially different administrative levels, adding an element of variation in the relationship between different cities and their rural populations. It must be noted these data sets conceal differences in data quality. One particular issue we will discuss

<table>
<thead>
<tr>
<th>Period</th>
<th>No. sites</th>
<th>Minimum site size</th>
<th>1st quartile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd quartile</th>
<th>St. dev.</th>
<th>Maximum site size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iron Age</td>
<td>549</td>
<td>0.01</td>
<td>0.1</td>
<td>0.2</td>
<td>3.72</td>
<td>1</td>
<td>17.02</td>
<td>185</td>
</tr>
<tr>
<td>Archaic</td>
<td>1,248</td>
<td>0.01</td>
<td>0.07</td>
<td>0.1</td>
<td>1.93</td>
<td>0.5</td>
<td>11.49</td>
<td>185</td>
</tr>
<tr>
<td>Post-Archaic</td>
<td>914</td>
<td>0.01</td>
<td>0.1</td>
<td>0.2</td>
<td>2.47</td>
<td>0.5</td>
<td>13.37</td>
<td>185</td>
</tr>
</tbody>
</table>

The data from the Tiber Valley Project have not been included (total number of sites: 539) because they do not provide the estimated size of sites.

**FIGURE 6** | Rank size in Etruria. (A) All sites in the Iron Age and Orientalizing. (B) All sites in the Archaic period. (C) All sites in the Post-Archaic Period. (D) All sites larger than 0.5 ha. in the Iron Age and Orientalizing. (E) All sites larger than 0.5 ha. in the Archaic period. (F) All sites larger than 0.5 ha. in the Post-Archaic period. The dark line represents the observed data. The grey lines represent the bootstrap results performed on random samples of sites.
later is that, whereas many of these territories have relatively good rural data sets and relatively good evidence for the size of the primate centre, the centres of intermediate size are not always captured within the survey area. Notably, we will return to the case of Chiusi later, where this problem is most critical.

**Settlement Hierarchy Derived From Rank-Size Analysis**

Table 1 provides a picture for Etruria of the central tendency and dispersion of settlements size (ha) in each period. We can see that the inter-quartile ranges (the 50% of values between the 3rd and the 1st quartiles, values that is between 0.07 and 0.5 ha) and the median between the Archaic and Post-Archaic periods differ only minimally, as confirmed by a Whitney–Wilcoxon test ($p = 0.95$). On the other hand, the interquartile range, and the mean between the Iron Age and the other two periods suggest general differences in settlement size distributions, as also confirmed by a Whitney–Wilcoxon test ($p < 0.01$). A notable difference in the Archaic period is the considerable increase in smaller rural settlement leading to the lower mean and a smaller standard deviation of site size. As a general trend, there was clearly more management of the countryside from within these smaller settlements.

Figures 6A–C shows rank-size analyses by using all the sites for each period. At a first glance, the size distributions appear similarly convex. The calculation of A-coefficients (Tables 2A–C) and the 95% confidence error ranges from the bootstrap method tell us that the rank-size curve is convex (Tables 2A–C). We can thus say confidently that the pattern is convex. As an extra precaution, since the rank-size in Figures 6A–C has a lower tail of very few settlements with size approximately equal to 0.1–0.01, the results might be distorted. Conventionally, urban geographers use the first 50/100 ranked sizes in a given area in order to avoid this “lower tail effect” (Hodder and Orton, 1976; Cristelli et al., 2012). For this reason, the analysis has been repeated for sites larger than 0.5 ha (Figures 6D–F). This provides better results and also avoids the distortion provided by a too skewed distribution arising out of very small sites. In this second scenario, the results show a more marked convex pattern (Tables 2D–F). On this basis, the results show convex curves in all the three periods. These results indicate that there was little political and economic integration among different competing city-states in Etruria between 1,000 and 350 BC. Literary evidence is not necessary to establish the differentiated political structure, since the convex curve of the rank-size provides convincing independent evidence on quantitative grounds.

After performing the above analysis on the whole of Etruria, we employ the XTENT model-defined territories in order to break down the study area into smaller political windows of analysis to investigate how settlement size distributions changed at the local scale. For this reason, we performed rank-size analyses for each XTENT model defined territories during the Iron Age/Orientalizing, Archaic and Post-Archaic periods. However, we were unable to run the analyses in all the territories because some of them do not contain sufficient data of both higher order settlement and rural settlement given the patchiness of the spatial coverage granted by the archaeological surveys carried out to date (or made available in open access, see Appendix A) in Etruria (see Figure 1B).

The rank-size analyses performed for the XTENT-defined territories during the Iron Age/Orientalizing period (ca. 1,020/1,000–580 BC) show strong primate patterns (Figure 7) and both the observed A-coefficients and the 95% confidence intervals show negative values (A2) exceeding −1, indicating that more than one observed settlements are smaller than the expected smallest settlement predicted by the Zip’s Law (see Table 3). Nevertheless, the results concerning Civita Castellana (Figure 7L) should be interpreted more cautiously given that their corresponding 95% confidence interval comprises also positive values which may indicate the possibility of a convex settlement pattern (see Table 3L).

The rank-size analyses performed for the XTENT-defined territories during the Archaic Period (ca. 580–480 BC) show similar strong primate patterns to the ones detected in the earlier periods (Figure 8 and Table 4). In this scenario, Fiesole (Figure 8B) and Acquarossa (Figure 8H) show patterns not as significantly primate, as suggested by the positive values in the 95% confidence interval (Tables 4B,H). Furthermore, the territory of Orvieto shows a prima-convex distribution (Figure 8I) with the overall A-coefficient (−0.51) resulting as the difference between the positive A1 values of the convex curve and the negative A2 values of the primate curve (Table 4I). The 95% confidence error range for A1 (0–0.37) and A2 (−0.22 to −0.95) indicates that the rank-size curve is significantly prima-convex (Table 4I). Finally, during the Post-Archaic period almost the totality of the XTENT defined territories show a strong primate pattern (see Figure 9 and Table 5) with the exception of Orvieto that is resilient with its prima-convex settlement system (Figure 9G and Table 5G).

Overall, it seems that in Etruria most XTENT-defined territories show a high settlement primacy from the early
Iron Age to the Hellenistic period, which is typical of city-states (see Figure 2). We can conclude that first millennium BC and pre-Roman Etruria was marked by a fragmented landscape of politically independent and competing polities, in which each exerted a strong centralized form of control within its own territory. It is also significant to note that even

FIGURE 7 | Rank size for the territories of Iron Age/Orientalizing centres with sufficient data points. (A) Volterra, (B) Populonia, (C) Murlo, (D) Vetulonia, (E) Vulci, (F) Bisenzio, (G) Acquarossa, (H) Orvieto, (I) Tarquinia, (J) Cerveteri, (K) Veio, (L) Civita Castellana. The dark line represents the observed data. The grey lines represent the bootstrap results performed on random samples of sites.
the smaller, “delicate” centres that existed in the interstices between the larger urban centres for the most part had a similar primate organization.

These results, taken at face value, whilst all primate, show the considerable variability in the relationship between different Etruscan cities and their countrysides. In early cultural research, scholars had pointed out the considerable variation in the cultural identity of Etruscan cities (Stoddart, 2020a). Many later scholars (Haynes, 2000; Pacciarelli, 2000; Bartoloni, 2003; Riva, 2010; Gliwitzky, 2015; Bell and Carpino, 2016; Naso, 2017; Shipley, 2017; Smith and Lulof, 2017) have built on the recognition of this fact, but it was Banti (1960) who was the first to express this effectively and succinctly. Earlier spatial work (Redhouse and Stoddart, 2011), using XTENT as outlined above, has concluded that the territorial size of each city varied considerably as a response to their geopolitical position. More recent research has shown the differing densities and disposition of medium-sized centres around the primate centres (Stoddart, 2016, 2020b). The systematic inclusion of rural settlement reinforces this interpretation. The density of rural settlement in the surrounds of the Etruscan cities varied from the high density in an area like Cerveteri to the very low density in an area like Perugia, with substantial variation in the spatially intervening centres.

The rank size data reveal similarly variable patterns. The main cities are substantially primate in their profile, whereas those on the margins and in the weaker buffer zones (e.g., Bisenzio and Civita Castellana in the Iron Age; Acquarossa in the Archaic; and Orvieto on the margins of the main Etruscan distribution in the Archaic and Post-Archaic period) tend to have a primo convex profile, suggesting a different relationship between centre and territory.

There is also a considerable range in the degree of centralization as measured by the A-coefficient. Veio was the most expansionist of urban centres, starting as a highly primate centre (−1.77; Table 3K) that extracted population from its hinterland (Ceccarelli and Stoddart, 2007), thereafter moving toward a less strong prymacy as its territory expanded (−0.73 and −0.58; Tables 4L, 5J). Cerveteri (−1.8; −1.08; −1.43; see Tables 3–5), its closest neighbour to the south east, had a territory hedged not only by the sea but by the powerful neighbours Tarquinia and Veio. This led to a notably more primate profile, where population was much more gathered within the precincts of the city and where perhaps the outlook was even more than Tarquinia toward the sea. This profile also fluctuated through time. The degree of centralization of buffer settlements, defined as those that do not survive the full sequence into the third phase (Civita Castellana, Bisenzio, Acquarossa, and Murlo), shows a much greater variability (−0.69 to −2.19), suggesting a wide gamut of strategies to survive in the political clutches of larger urban entities.

**Regional Centralization**

We measured local regional centralization for three urban centres providing robust settlement data from quite intensive archaeological surveys: Cerveteri, Murlo, and Tuscania. Figure 10 shows 10 concentric rings radiating out from these three urban centres and the graphed proportions of the total estimated settlement size within each ring (or “donut”). Cerveteri and Murlo show that most of the population concentrates in the innermost ring, while the successive rings (from the second onwards) show very low proportions of values (see Tables 6–10). Above all during the Iron Age and Post-Archaic period almost the total population is concentrated in the innermost ring (Tables 6, 8, 9). The high lines on the left side of the graph and the B-coefficient values ranging from 0.70 to 0.94 indicate that Murlo and Cerveteri exerted a high centralized control over the farming communities of their immediate surrounding hinterlands (Figures 10A,B). By contrast, Tuscania shows a more even distribution of the population across the 10 concentric rings. The lack of a very large urban centre superimposed on a tier of many smaller sites is indicated by very low B-coefficient values ranging from 0.06 to 0.24 (Figure 10C and Tables 12–14).

---

**TABLE 3** | A-coefficient values and bootstrapped error ranges for log scale rank-size curves of the XTENT-defined territories in Etruria in the Iron Age/Orientalizing.

<table>
<thead>
<tr>
<th>Letter in Figure 7</th>
<th>Territory</th>
<th>No. sites</th>
<th>Area Km sq.</th>
<th>Largest site (approx. ha)</th>
<th>Observed A-coefficient</th>
<th>Error range (95% confidence)</th>
<th>Curve shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Volterra</td>
<td>19</td>
<td>1.492</td>
<td>100</td>
<td>−1.92</td>
<td>2.42 (−3.04 to −0.62)</td>
<td>Primate</td>
</tr>
<tr>
<td>B</td>
<td>Populonia</td>
<td>38</td>
<td>972</td>
<td>150</td>
<td>−1.39</td>
<td>2.12 (−2.27 to −0.15)</td>
<td>Primate</td>
</tr>
<tr>
<td>C</td>
<td>Murlo</td>
<td>24</td>
<td>77.9</td>
<td>10</td>
<td>−2.19</td>
<td>2.40 (−3.21 to −0.81)</td>
<td>Primate</td>
</tr>
<tr>
<td>D</td>
<td>Vetulonia</td>
<td>35</td>
<td>1,051</td>
<td>100</td>
<td>−1.43</td>
<td>2.18 (−2.31 to −0.13)</td>
<td>Primate</td>
</tr>
<tr>
<td>E</td>
<td>Vulci</td>
<td>63</td>
<td>1,052</td>
<td>126</td>
<td>−1.50</td>
<td>2.04 (−2.36 to −0.32)</td>
<td>Primate</td>
</tr>
<tr>
<td>F</td>
<td>Bisenzio</td>
<td>17</td>
<td>408</td>
<td>35</td>
<td>−1.11</td>
<td>1.91 (−1.87 to 0.04)</td>
<td>Primate</td>
</tr>
<tr>
<td>G</td>
<td>Acquarossa</td>
<td>7</td>
<td>348</td>
<td>30</td>
<td>−1.29</td>
<td>2.44 (−2.54 to −0.10)</td>
<td>Primate</td>
</tr>
<tr>
<td>H</td>
<td>Orvieto</td>
<td>8</td>
<td>830</td>
<td>85</td>
<td>−1.77</td>
<td>2.43 (−2.90 to −0.47)</td>
<td>Primate</td>
</tr>
<tr>
<td>I</td>
<td>Tarquinia</td>
<td>107</td>
<td>959</td>
<td>150</td>
<td>−0.71</td>
<td>1.06 (−1.18 to −0.12)</td>
<td>Primate</td>
</tr>
<tr>
<td>J</td>
<td>Cerveteri</td>
<td>57</td>
<td>784</td>
<td>160</td>
<td>−1.80</td>
<td>2.40 (−2.77 to −0.37)</td>
<td>Primate</td>
</tr>
<tr>
<td>K</td>
<td>Veio</td>
<td>21</td>
<td>1,187</td>
<td>185</td>
<td>−1.77</td>
<td>2.12 (−2.81 to −0.69)</td>
<td>Primate</td>
</tr>
<tr>
<td>L</td>
<td>Civita Castellana</td>
<td>7</td>
<td>214</td>
<td>26</td>
<td>−0.69</td>
<td>2.0 (−1.92 to 0.22)</td>
<td>Primate or convex</td>
</tr>
</tbody>
</table>
FIGURE 8 | Rank size for the territories of Archaic centres with sufficient data points. (A) Volterra, (B) Fiesole, (C) Murlo, (D) Populonia, (E) Vetulonia, (F) Chiusi, (G) Vulci, (H) Acquarossa, (I) Orvieto, (J) Tarquinia, (K) Cerveteri, (L) Veio. The dark line represents the observed data. The grey lines represent the bootstrap results performed on random samples of sites.
TABLE 4 | A-coefficient values and bootstrapped error ranges for log scale rank-size curves of the XTENT-defined territories in Etruria in the Archaic period.

<table>
<thead>
<tr>
<th>Letter in Figure 8</th>
<th>Territory</th>
<th>No. sites</th>
<th>Area Km sq.</th>
<th>Largest site (approx. ha)</th>
<th>Observed A-coefficient</th>
<th>Error range (95% confidence)</th>
<th>Curve shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Volterra</td>
<td>68</td>
<td>4,201</td>
<td>100</td>
<td>−1.36</td>
<td>1.72 (−2.12 to −0.40)</td>
<td>Primate</td>
</tr>
<tr>
<td>B</td>
<td>Fiesole</td>
<td>23</td>
<td>1,200</td>
<td>30</td>
<td>−1.62</td>
<td>2.93 (−2.42 to −0.51)</td>
<td>Primate or convex</td>
</tr>
<tr>
<td>C</td>
<td>Murlo</td>
<td>71</td>
<td>329</td>
<td>10</td>
<td>−1.29</td>
<td>1.71 (−2.03 to −0.32)</td>
<td>Primate</td>
</tr>
<tr>
<td>D</td>
<td>Populonia</td>
<td>94</td>
<td>1,576</td>
<td>150</td>
<td>−1.30</td>
<td>1.80 (−2.10 to −0.30)</td>
<td>Primate</td>
</tr>
<tr>
<td>E</td>
<td>Vetulonia</td>
<td>57</td>
<td>2,174</td>
<td>100</td>
<td>−1.17</td>
<td>1.87 (−1.83 to 0.04)</td>
<td>Primate</td>
</tr>
<tr>
<td>F</td>
<td>Chiusi</td>
<td>23</td>
<td>1,544</td>
<td>50</td>
<td>−2.35</td>
<td>2.70 (−3.39 to −0.69)</td>
<td>Primate</td>
</tr>
<tr>
<td>G</td>
<td>Vulci</td>
<td>98</td>
<td>1,812</td>
<td>126</td>
<td>−1.13</td>
<td>1.66 (−1.86 to −0.20)</td>
<td>Primate or convex</td>
</tr>
<tr>
<td>H</td>
<td>Acquarossa</td>
<td>4</td>
<td>242</td>
<td>30</td>
<td>−0.92</td>
<td>2.04 (−1.84 to 0.20)</td>
<td>Primate or convex</td>
</tr>
<tr>
<td>I</td>
<td>Orvieto</td>
<td>25</td>
<td>2,086</td>
<td>85</td>
<td>−0.51</td>
<td>A₁ = 0.37 (0.37 to 0)</td>
<td>Primo-convex</td>
</tr>
<tr>
<td>J</td>
<td>Tarquinia</td>
<td>266</td>
<td>1,352</td>
<td>150</td>
<td>−0.44</td>
<td>0.90 (−0.81 to 0.09)</td>
<td>Primate</td>
</tr>
<tr>
<td>K</td>
<td>Cerveteri</td>
<td>295</td>
<td>801</td>
<td>160</td>
<td>−1.08</td>
<td>1.77 (−1.80 to −0.03)</td>
<td>Primate</td>
</tr>
<tr>
<td>L</td>
<td>Veio</td>
<td>119</td>
<td>1,862</td>
<td>185</td>
<td>−0.73</td>
<td>1.10 (−1.22 to −0.12)</td>
<td>Primate</td>
</tr>
</tbody>
</table>

DISCUSSION

Data Assessment and Self-Critique

As can be seen in Figure 1B, and briefly mentioned above, the availability of data varies across the territorial landscapes defined by spatial coverage and intensity of the archaeological surveys carried out in the area, providing a current limit to the historical validity of the data. By comparison with the classic studies of state formation in the arid areas of the Near East, the availability of data varies across the territorial landscapes. As can be seen in Figure 2B, C, the data is much more comprehensive because the important data infill for Tarquinia, Vulci, and Chiusi lies somewhere between these extremes. In all these cases the main focus of survey recovery is at the limits of the territory toward the tentatively defined frontier even though the work is generally of high quality (Tarquinia: Quilici Gigli, 1970; Vulci: Carandini et al., 2002; Chiusi: Paolucci and Francovich, 2007). A crucial addition has been the acquisition of the more recent survey of Tuscania, defined as resting in the territory of Tarquinia by XTENT (Barker and Rasmussen, 1988). Equally crucial is the availability of data from the important region of Veio, gathered for many years by the British School at Rome (Patterson et al., 2000a) and now re-analyzed (Patterson et al., 2020). This complements the later work of Rajala (2007, 2013) around Nepi. At an interregional scale (e.g., Judson and Hemphill, 1981; Guidi, 1985; Palmisano et al., 2017, 2018) large data sets may allow general patterns to be detected even if there is data loss at a local level. At a more local level, such as analysis of individual territories, the impact of data loss and spatial skewing needs to be taken into account since these factors can affect satisfactory interpretation. On a selective basis we now take steps to illustrate the detection and rectification of this loss of information.

It must be emphasized that these results at the regional level require considerable attention to detail in order to assess political reality. It has already been noted that the sampling for rural settlement is sometimes poorly located within the territory and that an unsupervised presentation of the results even of relatively large data sets can create unexpected results. One example is the profile of Chiusi. The results presented by the data purely from the XTENT data and the rural survey data led to the unexpected result discussed earlier that Chiusi is expressed as consistently the most primate territorial organization of Etruria during the Archaic and Post-Archaic periods (A coefficient = −2.35 and −1.85, Tables 4F, 5E). A closer examination reveals...
FIGURE 9 | Rank size for the territories of Post-Archaic centres with sufficient data points. (A) Volterra, (B) Fiesole, (C) Populonia, (D) Vetulonia, (E) Chiusi, (F) Vulci, (G) Orvieto, (H) Tarquinia, (I) Cerveteri, (J) Veio. The dark line represents the observed data. The grey lines represent the bootstrap results performed on random samples of sites.
TABLE 5 | A-coefficient values and bootstrapped error ranges for log scale rank-size curves of the XTENT-defined territories in Etruria in the Post-Archaic period.

<table>
<thead>
<tr>
<th>Letter in Figure 9</th>
<th>Territory</th>
<th>No. sites</th>
<th>Area Km sq.</th>
<th>Largest site (approx. ha)</th>
<th>Observed A-coefficient</th>
<th>Error range (95% confidence)</th>
<th>Curve shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Volterra</td>
<td>45</td>
<td>6,556</td>
<td>100</td>
<td>−1.34</td>
<td>1.64 (−2.08 to −0.44)</td>
<td>Primate</td>
</tr>
<tr>
<td>B</td>
<td>Fiesole</td>
<td>41</td>
<td>1,831</td>
<td>30</td>
<td>−1.33</td>
<td>1.98 (−2.05 to −0.07)</td>
<td>Primate</td>
</tr>
<tr>
<td>C</td>
<td>Populonia</td>
<td>91</td>
<td>2,169</td>
<td>150</td>
<td>−1.47</td>
<td>2.11 (−2.16 to −0.05)</td>
<td>Primate</td>
</tr>
<tr>
<td>D</td>
<td>Vetulonia</td>
<td>46</td>
<td>2,415</td>
<td>100</td>
<td>−1.32</td>
<td>1.91 (−1.86 to 0.05)</td>
<td>Primate</td>
</tr>
<tr>
<td>E</td>
<td>Chiusi</td>
<td>53</td>
<td>2049</td>
<td>50</td>
<td>−1.85</td>
<td>2.06 (−2.74 to −0.68)</td>
<td>Primate</td>
</tr>
<tr>
<td>F</td>
<td>Vulci</td>
<td>104</td>
<td>2,237</td>
<td>126</td>
<td>−0.92</td>
<td>1.27 (−1.48 to −0.21)</td>
<td>Primate</td>
</tr>
<tr>
<td>G</td>
<td>Orvieto</td>
<td>20</td>
<td>2,781</td>
<td>85</td>
<td>−0.49</td>
<td>A1 = 0.33 (0.33 to 0)</td>
<td>Primo-convex</td>
</tr>
<tr>
<td>H</td>
<td>Tarquinia</td>
<td>264</td>
<td>1,755</td>
<td>150</td>
<td>−0.35</td>
<td>0.68 (−0.65 to 0.03)</td>
<td>Primate</td>
</tr>
<tr>
<td>I</td>
<td>Cerveteri</td>
<td>65</td>
<td>640</td>
<td>160</td>
<td>−1.43</td>
<td>2.06 (−2.28 to −0.22)</td>
<td>Primate</td>
</tr>
<tr>
<td>J</td>
<td>Veio</td>
<td>172</td>
<td>2,622</td>
<td>185</td>
<td>−0.58</td>
<td>0.98 (−1.02 to −0.04)</td>
<td>Primate</td>
</tr>
</tbody>
</table>

FIGURE 10 | Graphs (left) and maps (right) showing the distribution of estimated settlements size across 10 concentric rings surrounding Cerveteri (A), Murlo (B), and Tuscania (C). The rings are ordered from innermost to outermost (left to right). Veio is not included in this analysis because of the absence of robust site size data. Murlo was not occupied in the Post-Archaic period.

that the data are composed of small rural settlements at the limits of the territory and that there was no inclusion of the intermediate size centres which have to be inferred from cemeteries ringing medieval and modern settlement closer to Chiusi (e.g., Castiglione del Lago, Città della Pieve, Cetona, Sarteano, Montepulciano, and Chianciano).

As an illustration of the rectification of this issue, new data for intermediate-sized settlement have been introduced for
two urban centres, Chiusi and Cerveteri, drawing on Stoddart (2020b) for informed guesses of some of the missing data. For Cerveteri, the work has drawn on Guidi (1985) and Judson and Hemphill (1981) to gauge the best response. For Chiusi, a judicious combination of Bianchi Bandinelli (1927) and Google Earth has been deployed to the area of likely settlement areas associated with known cemeteries. The reworked data are presented here for both the territories of Chiusi and Cerveteri in the Archaic period as an illustration of the diagnostic procedures required when employing large data sets at a local level. These results are shown spatially (Figure 11) and as simple size classes (Figure 12A) and as rank size (Figure 12B). Similar problems most probably affect other territorial profiles to a less severe degree, but the situation with Chiusi illustrates the problem most clearly since the interpretation of political power within the territory changes completely. Chiusi is now close to Zipfian lognormal (Table 14) suggesting a much more balanced degree of centralization compared with most of the other Etruscan centres. It also has a newly deciphered stepped hierarchy of settlement (Figure 12Ba), raising questions of the relationship between Zipfian patterns and Christaller’s (1933) ideal models. Cerveteri, by contrast, retains its status as a very powerful place distant in size from the next largest settlement (Figure 12Bb).

Other problems can arise from boundary problems and system closure. In circumstances where more than one settlement system is pulled into the same analysis, a convex distribution of sizes may well result. The multi-scalar approach adopted here with the recognition of all the more significant nucleations, both the primate and the second-order nucleations, with boundaries determined by the XTENT technique, goes some way toward overcoming this potential problem.

The statistical (bootstrapping) approach taken here also guards against problems of statistical uncertainty. Where the simulated examples show a wide range between positive and negative value of the coefficient A, we should be more cautious of the interpretation. We can note that our qualitative observations about the data for Tarquinia (good) and Chiusi (less good) are substantiated by the statistical observations on this basis.
In both the qualitative and quantitative assessments Tarquinia seems to be both statistically most robust and culturally clear. By contrast, Chiusi appears to be at the other extreme with the wide error range and to be counter intuitive to what might be expected culturally. As a variation on this theme, Civita Castellana, Fiesole, and Acquarossa show envelopes that cross the log normal boundary between convexity and primacy, suggesting another form of uncertainty (Tables 3L, 4B,H).

A further analysis was undertaken on the three most spatially robust data sets: Cerveteri (after the enhancement above), Tuscania (after the addition of the new data set), and Murlo. These are very different centres (as already discussed) but share similar qualities of data. Buffering of concentric areas moving away from the urban centre has been able to demonstrate the density of rural settlement (by number and area) as one moves away from the centre of urban power (Figure 10 and Tables 6–14). From this analysis, it appears that both the highly centralized primate centre (Cerveteri) and the opportunistic nucleated centre (Murlo), that ultimately failed to endure, were engaged in similarly strong strategies of centralization. By contrast, Tuscania, a subsidiary centre, was much less strongly nucleated.

More tentatively we can assess, in four cases, the distribution of the number of sites at increasing distance from the centre (Figures 13, 14). These latter results are affected by the survey sampling zones, but do reveal interesting differences in scale, clustering, and potential buffer zones, building on initial indications given by Rendeli (1993). These results are presented in two forms, the first (Figure 13) normalized on a scale from 0 to 1 in order to compare profiles independent of the very different sizes of the settlement in the three cases, the second (Figure 14) as raw densities, which we must caution may also be affected by the research intensity. The two primate centres show different profiles in the normalized results. Cerveteri presents a more concentrated profile of rural settlement (particularly before the Post-Archaic), perhaps constrained by its smaller terrestrial territory (Figure 13A). By contrast, Veio exhibits its expansionist territorial ambitions by a wider distribution of rural settlement (Figure 13D). The two smaller settlements, one subsidiary (Tuscania), the other “delicate” (Murlo) present a much more...
TABLE 14 | A-coefficient values and bootstrapped error ranges for log scale rank-size curves of the XTENT-defined territories for updated data for Chiusi and Cerveteri in the Archaic period. See Figure 12B for reference.

<table>
<thead>
<tr>
<th>Territory</th>
<th>No. sites</th>
<th>Area Km sq.</th>
<th>Largest site (approx. ha)</th>
<th>Observed A-coefficient</th>
<th>Error range (95% confidence)</th>
<th>Curve shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiusi</td>
<td>31</td>
<td>1,544</td>
<td>50</td>
<td>−0.77 (−0.02 to −0.79) (A₁−A₂)</td>
<td>1.55 (−1.32 to 0.23)</td>
<td>Zipfian</td>
</tr>
<tr>
<td>Cerveteri</td>
<td>299</td>
<td>801</td>
<td>160</td>
<td>−0.87</td>
<td>1.25 (−1.40 to −0.15)</td>
<td>Primate</td>
</tr>
</tbody>
</table>

FIGURE 11 | Presentation of new data for Chiusi (f) and Cerveteri (k) in spatial form.

locally clustered profile, particularly if one bears in mind that Murlo itself did not exist in the Post-Archaic (Figures 13B,C). This profiling is supplemented by the evidence for the absolute density of settlement around these centres (Figure 14). These show the greatest density around the subsidiary centre of Tuscania, followed in turn by Cerveteri and Murlo, whereas Veio generally has the lightest density, compensated by its greater extension of higher density away from the urban centre. However, this picture is biased by the research intensity of the archaeological surveys carried out in the region. Unlike in the case of the other three centres, the area around Tuscania was intensively surveyed by using field-walking transects (cf. Barker and Rasmussen, 1988). Cerveteri and Veio have a halo of low density of rural settlement close to the urban centres perhaps partly occupied by cemeteries. In the case of Cerveteri, this takes the form of a denser band in the Archaic period at a distance between 3 and 6 km from the urban centre (Figures 14A,D).

Although not covered in these diagrams, this is also the place to note the very low levels of density of Etruscan rural settlement in intensively surveyed areas such as the Cecina valley, at the limits of some northern Etruscan urban territories.

Finally, we performed a multi-scalar spatial statistics technique known as Ripley’s K function to assess if the major 25 urban centres of Etruria were spatially clustered or segregated (see Ripley, 1976). We used Monte Carlo simulations of random point distributions to build a 95 per cent confidence envelope of the null hypothesis of complete spatial randomness (CSR) by carrying out 1,000 iterations (Bevan and Conolly, 2006, p. 220; Palmisano, 2013, p. 351). These estimates were then compared with the observed values of K (L) in order to obtain a statistically robust measure of a clustered or even point distribution in our study area (Figure 15). Sites are clustered when the solid line is above the grey envelope and evenly distributed when the line is below the grey envelope. From this analysis, it
FIGURE 12 | (A) Presentation of new data for Chiusi (a) and Cerveteri (b) as histograms of site size frequency. (B) Presentation of new data for Chiusi (a) and Cerveteri (b) as Rank Size.

can be seen that the major centres are evenly distributed at a radial distance between 15 and 30 km. This would confirm the idealized pictures provided about the radius of 15–30 km of city-states (e.g., Hansen, 2000) and also informs us that the major centres shared the space evenly and established well-defined individual catchments, probably on the basis of land holdings, subsistence strategies, and local political identity. This analysis shows the general trends of Etruria at a supra-regional scale compared with the detailed variation shown by XTENT at the local/regional scale.

From Large Data to Urban Anthropology

These results point to the general processes of expanding territories from their centres and yet multiple local origins and outcomes in the development of Etruscan urbanism (Stoddart, 2018). We know enough of the economic infrastructure to establish the importance of a number of common factors. The first is the Mediterranean polyculture of cereals, vine and olive. The annual cycle of grain agriculture had been present for five millennia, but the longer term care of tree crops necessitated a protection of the landscape that tied in well with urbanism in more ways than simply providing a potable component for the feast (Stoddart et al., 2019). It required protection of a perennial resource close to the urban centre. The landscape was not opened up as much as in the Roman period, but pollen studies show that, at least at a local level, the vegetation was manipulated not only to provide cereal and tree crops, but also grazing for sheep, cattle, and pigs (Stoddart et al., 2019). In the area around the city of Veio, watercourses were manipulated to ensure the provision of sufficient water for crops, and many cities show the construction of wells and drainage systems to maintain a high living standard (Judson and Kahane, 1963). Some cities such as Veio in the south and probably Chiusi in the north specialized in agricultural production. Other cities, such as Populonia in the north, located close to the ore-bearing hills ranging toward Elba, specialized in iron production. Further cities, notably the coastal city of Cerveteri to the south, focused on trading activity. These essential developments enabled the foundation and maintenance of the stable nucleated centres, lasting very many generations, in contrast with a number of other contemporaneous civilizations, and yet setting up essential variations between them in the organization of their territories.

As we have seen from the XTENT analysis, there was one Etruscan city which could have created a different classical history: Veio, the most southerly Etruscan city. If it had defeated Rome, it is conceivable that a Veian empire would have come down to us today as the contemporary rival to the Han Empire of China, not the Roman empire every European citizen knows. If history had been different, we would have had an Etruscan historiography, that no doubt would have made disparaging comments about the defeated Romans, that would have given us a detailed mythological genealogy of Etruscan city foundations similar to Romulus and Remus, that would have outlined in some detail the great families through time (the descent groups of anthropology) and the politics of the time. It is no accident that these two rival cities, Veio and Rome, Etruscan and Latin, faced each other across the Tiber, displaced from the spatial centre of their respective cultures. Since each was dominant in its own political world, one was likely to prevail.

At the risk of social determinism, there was little risk of such an Etruscan empire. The reason lies in another account of Etruscan origins. This account derives from the specific nature of the political decision to move from a village society to one of nucleation or urbanism. This decision took place almost simultaneously between 1,000 and 900 BC in all five major cities—Veio, Orvieto, Cerveteri, Tarquinia, and Vulci—of Southern Etruria. We have seen it played out in the political landscape discussed above. The outcome was a relative equilibrium of like-sized and like-politicized communities (Figures 2, 15), where an internal tension existed between descent groups and community and an external tension between the urban communities themselves. But for the presence of Rome, Veio might have surfaced pre-eminent from this dynamic equilibrium; however the sack of Veio in
FIGURE 13 | Normalized sites density at increasing distance (concentric rings) from Cerveteri (A), Murlo (B), Tuscania (C), and Veio (D). Murlo was not occupied in the Post-Archaic period.

396 BC (according to the Romans) truncated this potential. The more probable outcome, even in the absence of Rome, would have been a long-standing rivalry between equally ranked cities, where no individual city was able to overcome its equally large neighbour. This is a pattern imprinted on the political landscape, as we have seen earlier in the paper (compare also discussion in Fulminante and Stoddart, 2013; Fulminante et al., 2017; Prignano et al., 2019; Fulminante, in press).

Some Etruscan origins were, however, unsuccessful, even before the intervention of Rome. The powerful stable centres outlined above had empty frontier zones at the edges of their territories that gave opportunity to other political origins. In these zones, particularly during the Orientalizing period, independently minded groups experimented with their own political organization, to varying degrees of success. This dynamic pattern had substantial similarity to the patterns of state formation noted in Africa by Kopytoff (1989), a process that he defined as the internal frontier. The dramatic image of a weather map where highs are surrounded by troughs offered opportunities for the discontented, those inflicted with witchcraft, to develop their own independent political opportunism. The political landscape of Etruria was anchored to the powerful places, but, with varying degrees of success, other centres attempted to forge their own political presence. Three zones can be identified where origins were failures at different scales. In the south, two relatively large centres, Bisenzio and Acquarossa, sheltered by a series of volcanic lakes, managed to maintain themselves for several centuries before they were squeezed out of existence by their even more powerful neighbours: Veio, Cerveteri, Orvieto, and Tarquinia. At the juncture between North and South Etruria, the Albegna valley system beyond the easy reach of Vetulonia to the north and Vulci to the south gave opportunity to two centres which followed one another in turn: Marsiliana in the eighth and seventh century and Doganella in the sixth century BC. Their origins were too unstable to prevail against the greater power of their neighbours or against Rome. Finally, in the Chianti zone of North Etruria, Murlo, and Castelnuovo Berardenga were small nucleated centres of great flamboyance, but destroyed, in the case of Murlo, on at least two occasions. In this latter case the origins of the Etruscans in the interstices between more powerful cities was fleeting and temporary.

As viewed from a perspective of modern social anthropology, origins are related to the construction of an identity at any
given moment in time. What the Etruscans themselves saw as their own origins is one substantive answer to the question of where they came from. An understanding of this approach is restricted by the lack of their own literature to frame exactly their response, although many constructive attempts have been made from their visual culture. What archaeologists can reconstruct as an economic and political sequence is another valid approach and this is what this paper has achieved for the political landscape. Here scholars are on much stronger ground as new research on the individual cities and territories has become more sophisticated, filling previous gaps in research. A synthesis can draw on both approaches, since they form two perspectives of the same picture. Thus, it is clear that every city had its own individual identity, differing in economic specialization, funerary practice, degree of centralization, and territorial control. More tentatively from sources that include iconography, we can establish that each of these cities had their own mythological account, sometimes drawing on local legitimization, in the form of figures such as Tarchon for Tarquinia, sometimes drawing on exotic ambiguous figures such as Ajax, whose exploits Etruscans claimed from an external world. Language and, to a certain extent ritual, point to a wider cultural unity, but in a secondary sense. Late Etruscan accounts project a suspiciously detailed self-professed understanding of their construction of time that ranged from the organization of 8-day weeks into 8-month years and 10 saecula that started in 1,201 BC and continued until 83 BC (following the scheme of Pfiffig, 1975). Such detail was only possible, once they were part of the Roman world which offered their elite better opportunities.

CONCLUSION

This paper has combined multiple approaches of spatial analysis to investigate patterns of urbanism focussed in a sub-region, Etruria, and a sub-period, the central section of the first millennium BC. The approach we have employed is an amplification, improvement, and complementary extension of past work (Stoddart, 1987, 2016, 2020a,b). The innovation here has been to use a large dataset of rural settlement (Table 1) from Palmisano et al. (2017, 2018) and four further unpublished data sets to explore the spatial characteristics of Etruscan urbanism in a more comprehensive, systematic, quantitative, and questioning way than has been possible hitherto, to add to the focus of previous authors (e.g., Rendeli, 1993) on urbanism processes as manifested by the evidently powerful places of urban inhabitation. We can write about many origins of the Etruscans.
Each of the large urban communities clearly had a profound sense of its origin which was also politically orchestrated through religion and sense of place, and space. There were many powerful places and thus many origins, some of which emasculated the origins and indeed existence of smaller places. The greatest place was ultimately and historically Rome, which produced its own annalistic account of its own and others’ origins. These others included the Etruscans who were classified by the Romans as military rivals, pre-occupied with religion, overly fat, and characterised by powerful women. Nevertheless, by concentrating on the landscape dimension of Etruscan urbanism, we hope that this study can also contribute usefully to wider discussions of the archaeology of early urbanism, moving beyond the “Mediterranean myopia” that some New World Scholars (e.g., Blanton, 2001) have suggested to be present in the local tradition of research.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the Appendix A of the paper.

REFERENCES


AUTHOR CONTRIBUTIONS

SS has taken the lead in proposing the project and synthesizing the results. AP whose previous work inspired the paper, provided the main body of rural data from his recent studies (Palmisano et al., 2017, 2018) within the territories defined by XTENT. He also executed the presentational statistics, rank size analyses, statistical assessments, and verification. DR previously worked with SS in completing a more sophisticated DEM-sensitive implementation of XTENT for the Etruscan area than the ones implemented previously by Harrison and Stoddart. The other co-authors have kindly contributed data on the Cecina Valley, Populonia, Tuscania, and Veio, substantially swelling the data set with systematically collected data. All authors contributed to the structure and argument of the paper.

ACKNOWLEDGMENTS

The publication of this paper was funded via an OPEN-AIRE grant to Dr. Francesca Fulminante through the University of Rome (Marie Curie Sklodowska Fellowship 628818- Past-People-Nets 2014–2016).


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

---

Copyright © 2020 Stoddart, Palmisano, Redhouse, Barker, di Paola, Motta, Rasmussen, Samuels and Witcher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
APPENDIX A: REPRODUCIBILITY AND SUPPLEMENTARY DATA

The raw dataset and the supplementary materials used in the present article have been deposited on the free accessible online repository Zenodo: https://zenodo.org/record/3735665. The digital archive related to this paper provides reproducible analysis in the form of six scripts written in R statistical computing language. The present repository contains also a R Markdown tutorial to drive step by step any practitioner interested in running some analytical tools to assess regional centralization and settlement hierarchies: site-size histograms, rank-size graphs, A-coefficient, and B-coefficient.
## Advantages of publishing in Frontiers

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPEN ACCESS</td>
<td>Articles are free to read for greatest visibility and readership</td>
</tr>
<tr>
<td>FAST PUBLICATION</td>
<td>Around 90 days from submission to decision</td>
</tr>
<tr>
<td>HIGH QUALITY PEER-REVIEW</td>
<td>Rigorous, collaborative, and constructive peer-review</td>
</tr>
<tr>
<td>TRANSPARENT PEER-REVIEW</td>
<td>Editors and reviewers acknowledged by name on published articles</td>
</tr>
<tr>
<td>REPRODUCIBILITY OF RESEARCH</td>
<td>Support open data and methods to enhance research reproducibility</td>
</tr>
<tr>
<td>DIGITAL PUBLISHING</td>
<td>Articles designed for optimal readership across devices</td>
</tr>
<tr>
<td>IMPACT METRICS</td>
<td>Advanced article metrics track visibility across digital media</td>
</tr>
<tr>
<td>EXTENSIVE PROMOTION</td>
<td>Marketing and promotion of impactful research</td>
</tr>
<tr>
<td>LOOP RESEARCH NETWORK</td>
<td>Our network increases your article’s readership</td>
</tr>
</tbody>
</table>

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org  | +41 21 510 17 00

FOLLOW US
@frontiersin