Introduction

Economic geographers seeking to understand how substantial variations in population concentration and economic activity are created across the landscape correctly note that, except in cases of colonial imposition, such variations are always the result of cumulative processes whereby initial natural advantages of particular sites or areas are extended and compounded by socially created technologies and institutions delivering increasing returns to scale. In this manner, they argue, self-reinforcing processes of accumulation, exchange, agglomeration, and innovation are created that ultimately determine the varying developmental trajectories of different regions and the location, number, and rate of growth of cities within them (Krugman 1991, 1995, 1998; Pred 1966).

The economist Paul Krugman (1996a) vividly illustrates this process in reference to the process of expansion of Chicago in the 19th century, as outlined by the historian William Cronon (1991) in his book *Nature’s Metropolis: Chicago and the Great West*. Cronon insightfully distinguishes between two settings in which the evolution of Chicago took place. The first was its “natural landscape,” entirely determined by geography and environment. The second was what he terms the “created landscape,” which resulted from human innovations and institutions that substantially altered and reshaped Chicago’s natural setting and significantly expanded the advantages of the city’s initial location for human settlement. Cronon argues that in the modern world the created landscape has become more important than the natural landscape as a determinant of urban location and regional developmental rates. Specifically, he sees Chicago’s initial role as a Great Lakes port, a role entirely determined by geography of the Great Lakes area, as eventually overshadowed by its later role as a railroad hub, a secondary but economically more important role that emerged as part of the “created landscape.” Chicago became the early economic center of its region because it was a port. Railroads later used Chicago as a hub precisely because it already was the early economic center of its region, and thereby helped make its initial centrality that much greater.

The insights of Cronon and Krugman about the ways in which natural and created landscapes determine, reinforce, and compound each other in modern cities and their surrounding areas are perhaps applicable to earlier cases of urban transformation. A case in point appears to be that of the emergence of early Sumerian civilization, widely acknowledged as the world’s earliest, along the alluvial lowlands of the Tigris–Euphrates rivers in what is today southern Iraq (southern Mesopotamia) during the second half of the fourth millennium BC. As Tony Wilkinson (2001) and Joan Oates (2001) have recently noted, this emergence took place after centuries, if not millennia, in which the developmental trajectory of polities in the southern Mesopotamian alluvium had hardly differed from that of neighboring societies across the ancient Near East. This becomes clear when we compare data pertinent for the fifth and fourth millennia produced by disparate surveys and excavations across north and southern Mesopotamia, southwestern Iran, and the Levant.

Briefly summarized, these data indicate that during the second half of the fifth millennium, Late Ubaid settlements in southern Mesopotamia (Oates 1983) were entirely
comparable in terms of both scale (roughly measured by settlement extent) and level of intra-site differentiation to contemporary (MS 3-Late Susiana) societies in the Susiana plain of Khuzestan (Delougaz and Kantor 1996; Wright 1984, Wright and Johnson 1975), and appear to have also been similar in scale to contemporary settlements in the Upper Euphrates, Upper Khabur, and Upper Tigris basins of Upper Mesopotamia (Kouchoukos and Hole 2003; Wilkinson 2000, 2003). Additionally, again, at least in terms of scale, Late Ubaid settlements in southern Iraq appear similar to contemporary Ghassulian Phase Chalcolithic settlements in portions of the Jordan Valley (Bourke 2001: 111–116).

A degree of differentiation in regional developmental rates starts to become apparent in some portions of the ancient Near East roughly at the transition from the fifth to the fourth millennia. The Levant drops out of the picture at this time as a result of a still not well understood process of collapse that marks a clear hiatus in indigenous processes of social evolution in that area (Levy 1998: 241–43). Nonetheless, development continued apace over large portions of “Greater Mesopotamia,” where “proto-urban” polities of considerable extent and complexity started to emerge during the early centuries of the fourth millennium BC, principally in portions of the Upper Khabur plains of northern Syria and northern Iraq, within the alluvial lowlands of the Tigris and Euphrates rivers in southern Iraq, and, to a lesser extent, in Susiana (Adams 1981; Nissen 1993; Kouchoukos and Hole 2003; Wilkinson 2000; 2003; Oates 2001).

Yet, only a few centuries later, by the second half of the fourth millennium, Upper and southern Mesopotamia were no longer developing largely in tandem or at comparable rates. Available evidence is clear, and shows that about the middle of the fourth millennium or so polities in the Tigris–Euphrates alluvial delta started to surpass their immediate neighbors and potential competitors across the Near East in terms of scale, degree of internal differentiation, and in the degree of hierarchy present in surrounding settlement grids. At this time, southern Mesopotamia became a dynamic regional cauldron comprised of multiple thriving and competing city-states, which formed part of a politically balkanized but culturally homogeneous early Sumerian civilization that extended at this point into southwestern Iran (but see Wright 1998 for a contrary opinion).

In contrast, by the third quarter of the fourth millennium, the early proto-urban sites of Upper Mesopotamia were in decline (Emberling 2002), just as a small number of colonies of southern Mesopotamian settlers were being established at strategic locations across the northern plains (Algaze 1993, 2001b, 2005; Schwartz 2001; Stein 1999, 2001). While indigenous societies continued to flourish across the north (Frangipane 2002), including in areas surrounding the intrusive southern settlements, as a group, Late Chalcolithic polities of Upper Mesopotamia at this time were no longer comparable in either scale or complexity to the, by now, much more developed polities of southern Mesopotamia, where a veritable revolution in human spatial, social, political, and economic organization had taken place.

The growing dichotomy in the historical trajectories of southern and Upper Mesopotamia reached a climax by the last century of the fourth millennium, when most of the Late Chalcolithic centers that existed in the northern plains had either severely retrenched or had been abandoned altogether. Indeed, available settlement pattern data suggests that by the start of the third millennium (early Ninevite V period) Upper Mesopotamia had been transformed into an overwhelmingly rural landscape characterized by multiple dispersed villages, and urbanism would not again arise in the north for at least
400 years (Algaze 1999: Table 2, with references). In contrast, the increasingly urbanized fourth millennium societies of the southern Mesopotamian alluvium continued to spiral upwards in scale, density, and complexity (internal differentiation) into the third millennium without any significant interruption (Adams 1981).

Sumerian civilization thus represents a dramatic “takeoff” – a decisive shift in favor of southern Mesopotamia of the balance of urbanization, socio-political complexity, and economic differentiation that had existed across the ancient Near East until the onset of the fourth millennium. Why did this shift take place? Could a comparable shift have occurred anywhere in the ancient Near East, or were there factors specific only to southern Mesopotamia that made it more probable that the shift would occur there rather than elsewhere? If the latter, what processes help account for the emergence of civilization in the south? And, finally, why did this emergence take place when it did, in the second half of the fourth millennium, and not before?

This paper is a preliminary attempt to explore these various questions relating to the Sumerian Takeoff. To be sure, as Lamberg-Karlovsky (1995, 2001) and others (e.g., Collins 2000) have repeatedly warned us, processes of such complexity cannot be fully explained without reference to concurrent ideological transformations, including new understandings about the relationship between the rulers and the ruled and new understandings about property and property rights (North and Thomas 1973). For this reason, ideological factors are often, and correctly, seen as having a central role in structuring early cities and states across the world (e.g., Wheatley 1971; Kolata 1983; Marcus 1983; Cowgill 2000). However, documenting the weight of ideological factors in the crystallization of pristine civilizations is inherently difficult because of the nature of the evidence at our disposal (below). Equally difficult is documenting transformations in how property was conceptualized in such societies, because a full written record is commonly absent in their initial phases. For this reason, the focus here is both narrower and simpler. I seek to elucidate the economic variables underlying the processes of urban growth and socio-economic differentiation in Middle and Late Uruk period southern Mesopotamia and, in so doing, to understand why developmental processes of comparable scale and dynamism were absent in neighboring societies at the time.

Specifically, I argue that the Sumerian Takeoff was the result, in part, of evolving, long-term trade patterns that differentially favored the development of societies in the alluvial lowlands of Mesopotamia over polities in neighboring regions. At first, the trade was spurred by natural differences in productivity between varying areas of the southern Mesopotamian alluvium and between the alluvium as a whole and surrounding areas–what Cronon refers to as the “natural landscape.” However, as the exchange unfolded over time, and as its scale and external scope increased, multiplier effects of the trade brought about substantial increases in the density and propinquity of populations in the alluvium as compared to those prevalent in competing areas at the time. This allowed for important social transformations to take place in southern Mesopotamia that, in turn, fall entirely in the realm of Cronon’s “created landscape.” Most important among these were (1) new forms of labor organization delivering economies of scale in the production of subsistence and industrial commodities and (2) new forms of record keeping that were much more capable of conveying information across time and space than the simpler reckoning systems used by contemporary polities elsewhere. These innovations furnished early
Sumerian polities of the fourth millennium with what turned to be their most important competitive advantage over neighboring societies.

In the sections that follow, I discuss the evidence that underpins these various contentions.

The Need for Models

Existing evidence for the emergence and growth of early cities in the alluvial environment of southern Mesopotamia throughout the various phases of the Uruk period (ca. 3900–3200/3100 BC) is of varying reliability, resolution, and coverage. The formative phases of the process remain shrouded in the mist of the so-called Early Uruk period (ca. 3900–3600 BC), a phase that for all practical purposes is known only through survey data (Nissen 1993). This immediately presents us with a significant obstacle to interpretation because by their very nature long-term historical processes can only be studied diachronically. Later phases of the Uruk period, the Middle (ca. 3600-3400/3300 BC) and Late Uruk (ca. 3400/3300-3200/3100BC) phases, are better understood, since pertinent data are provided by settlement pattern surveys, excavations at a small number of sites, a fairly extensive corpus of iconographic representations, and by a small amount of textual documentation. However, even with this extended evidentiary base there are still problems. Although existing excavations can be hugely informative for individual sites such as Uruk–Warka, Warka is almost certainly not representative for the alluvium as a whole, where the number of excavated Uruk sites remains small and the extent of exposures at those sites smaller still. Moreover, even at Warka, existing exposures remain unrepresentative outside of the intensively studied Eanna and Anu Precincts, and are pertinent mostly to the final phase of the Uruk period (Warka, Eanna IV; see Nissen 2001, 2002). Texts, likewise, exist only in the final stage of the period, and they are generally not particularly informative with respect to many questions for which we need answers. Moreover, in many cases they remain quite difficult to interpret (Englund 1998). Pictorial representations in a variety of media offer an entry into the social ideologies and cosmography of the Uruk world, and can be quite informative about the ritual and mundane activities of elite individuals depicted repeatedly (Schmandt-Besserat 1993) but, again, these images fail to illuminate a representative cross-section of early Sumerian society.

Since existing archaeological, textual, and art historical data from fourth millennium Mesopotamia are insufficient by themselves to clarify the processes of emergence and growth of early cities and states in the area, this paper invokes the work of economic geographers and developmental economists who seek to understand why, where, and how cities emerge in the modern world in order to generate testable propositions that will help us to better interpret what evidence we do have from Uruk Period Mesopotamia. These scholars commonly take it as a given that trade is a key factor in the evolution of social complexity and that cities serve as the cross-culturally most efficient way to manage regional and inter-regional exchange in situations marked by asymmetries in resource endowments, commodity production, and access to transportation across the landscape (e.g., Hicks 1969; O’Sullivan 1996; Fujita, Krugman, and Venables 1999). Economic historians and, to a lesser degree, some archaeologists, too, have long argued for the primacy of trade in explaining, for instance, phenomena as diverse as the establishment of colonial cities across the Mediterranean coast of Europe in the Iron Age (Wells 1980) and
the growth of urban centers across Europe throughout the Medieval periods (e.g., Pirenne 1936; Fox 1971, 1991; McCormick 2001).

In contrast, discussions in ancient Near Eastern studies about the why, where, and how big of early urbanism have, in my opinion, paid insufficient attention to the work of classical economists regarding the social ramifications of trade and the conditions that give rise to it. While pertinent work on the effects of population growth and the impact of regional and interregional trade did figure in some attempts to understand the formation of early cities and states in the ancient early Near East written during the 1970’s (for instance, Adams 1981; Service 1975; Smith and Young 1972; Young 1972; and Wright and Johnson 1975), these subjects have received comparatively little attention of late. Moreover, other topics of equal importance for understanding early urban processes have never been given the weight they deserve by scholars of the ancient Near East. Prominent among these are studies of how geographically determined differences in resource endowments and access to transport may have contributed the formation of early cities in the area, and how differences in technologies of communication may have contributed to the further growth of such cities after they crystallized. This is surprising because resource variability has been recognized as central to spurring economic activity at least since the work of David Ricardo (1817 [1971]), and technologies of transportation and communication have often been identified as crucial to processes of social evolution by scholars in various disciplines (e.g., Bairoch 1988; Fox 1980; McNeill 2000; Shennan 1999; Spufford 2002; Vance, Jr. 1986). Worse still is the fact that, with the notable exception of Hans Nissen (1976, 2000), ancient Near Eastern scholars have yet to fully address the role that organizational efficiencies yielding significant economies of scale (increasing returns) in the production and distribution of commodities (Krugman 1991) may have had in the development of early Mesopotamian cities. The continuing paucity of such research is particularly regrettable because economists have long known that the adoption of practices that systematically lower transaction costs and promote high rates of innovation can, like differences in resource endowments and access to transportation, be a critical factor in the emergence of sharply unequal developmental rates between regions, and in determining how long such divergent rates of growth can be sustained (North 1991).

In great part the lack of an economic focus of research into ancient Near Eastern urban origins is a direct consequence of the nature of the archaeological and textual data at hand for early cities in the area, which is difficult to characterize and impossible to quantify. However, several additional contributing factors, both conceptual and methodological, also exist.

On the conceptual front, one factor was that, in rejecting the well-documented excesses of early 20th century diffusionism, many archaeologists strove instead to explain past cultural changes too narrowly as mere adaptive reactions to ecological transformations. In so doing, their work fails to give proper consideration to a slew of other potential factors, which are also crucial to the explanation of pre and proto-historic social change, including long distance exchange and the cross-cultural interactions engendered by that exchange. The same was often the case at the opposite end of the ideological spectrum, as many Marxist archaeologists placed undue emphasis on changes in how commodities were produced in their analyses of ancient social transformations to the detriment of analyses focusing on the role of cross-cultural trade as both trigger and ongoing spur for those transformations.
Also problematic is the implicit assumption shared by many in our field that ancient socio-economic phenomena were of an essentially different nature from modern ones. This position considers the wealth maximizing behavioral postulate typically propounded by economists as entirely inappropriate for the study of both non-western (Sahlins 1972) and pre-modern societies (Polanyi 1957, 1977; Finley 1985). A logical corollary of this position is that the forces underlying the emergence and growth of early cities in antiquity must have also been different from those at work in historical periods closer to us.

But is this really the case? I suspect that the answer is yes in some cases and no in others, and that this latter answer would be quite clear if only we had the ability to precisely quantify and accurately characterize the types of economic activity in many urban societies of the pre-modern era. To be sure, Polanyi's admonition to the effect that economic behavior in early pre-industrial societies is always embedded in broader sociopolitical and ideological systems is patently correct, and so too is Finley's (1985) insistence on the centrality of status as a determinant of economic interactions in antiquity (on this point, see also Helms 1988, 1993). However, in spite of these peculiarities, substantial continuities in economic processes still exist across very different historical eras (North 1977: 709; Silver 1995; Shennan 1999; Snooks 1993) – at least since the rise of the earliest cities and states.

Early Mesopotamian civilization is a case in point. Polanyi's insistence, for example, on the absence of price-making markets in early Mesopotamia subject to laws of supply and demand has now been convincingly refuted by work that shows how exchange ratios for different commodities in some Mesopotamian cities varied according to changes in supply resulting from the ebb and flow of political and military circumstances during the Old Babylonian period, the first quarter of the second millennium BC (Farber 1978; Silver 1995. See also Warburton 2003). Equally relevant is recent work by Assyriologists clarifying the entrepreneurial and profit-seeking nature of contemporary Mesopotamian trade patterns, including Old Assyrian trade with Anatolia (Adams 1974; Dercksen [ed.] 1996, Larsen 1976) and Old Babylonian trade in the Persian Gulf (Oppenheim 1954; Leemans 1960).

While the degree of private risk-taking displayed in these texts is unusual for Mesopotamian economic forms predating the second millennium, there is no shortage of evidence for wealth-maximizing behaviors in the context of earlier Mesopotamian institutions and groups. Ancient Near Eastern scholars now broadly agree that in many cases the very same institutions that were at the center of early Mesopotamian centrally managed economies during the Early Dynastic Period also engaged in what can only be described as wealth-maximizing behaviors incompatible with Polanyi's characterization of those economies. Most commonly this took the form of urban temples of mid-third millennium date using silver derived from the conversion of accumulated agricultural surpluses into usable capital (Powell 1996) to finance risky trading ventures in order to acquire non-local resources for profit, whether alone or in conjunction with palaces (Leemans 1960; Postgate 1972). Such ventures were led by merchants who, at times, were clearly subordinates of the sponsoring institutions and, at times, appear to have been wholly or partially independent of them (Powell 1977). Given this evidence, why should we believe that comparable economic behaviors and motivations did not, and could not,
exist earlier, at the time of the inception of Mesopotamian urbanism in the fourth millennium?

Methodological factors also contributed to the general lack of an economic focus of much research into ancient Near Eastern urban origins. Central Place Theory, the tool most commonly used in approaching this subject, is inherently ill equipped to address questions of origin. In essence, the locational theories put forth by 20th century successors of von Thünen, most notably Christaller (1933 [1966]) and Lösch (1940 [1954]), seek to understand the forces that spread economic activity away from a center as a result of the tradeoff between economies of scale, which provide an incentive to concentrate production, and transportation costs, which provide an incentive to disperse production and managerial functions to multiple sites close to consumers/workers. In elucidating the interplay between centripetal and centrifugal economic forces determining a region’s settlement structure, Central Place models help us to understand how spatial hierarchies of function are maintained across a landscape. However, because the models simply assume the a-priori existence of a capital administrative and market center, they tell us nothing about why population and economic activity become concentrated in the first place, as John Marshall (1989) and Paul Krugman (1995) have insightfully noted.

Urban Dynamics: Why, Where, and How?

To understand the deeper causes of why and where cities emerge, we need to go back to the concept of comparative advantage, articulated by the great economist Ricardo almost two centuries ago and expanded upon by many economic geographers since then. From a modern Ricardian perspective, cities invariably represent nodes in wider transportation networks and serve to mediate trade between regions and polities with varying degrees of competitive advantage in the production of both necessary and desirable resources. Such advantage is created by differences in productivity between polities caused by a combination of the naturally uneven distribution of resource endowments in different areas, as Ricardo himself noted, differences in access to and in the cost of transportation, and social factors, such as organizational and technological efficiencies delivering increasing returns to scale, as noted by many of Ricardo’s intellectual successors (Balinger 2001; Krugman 1991, 1995; Fujita, Krugman, and Venables 1999).

Since cities form as a response to regional imbalances in productive advantages, they should preferentially emerge at natural passage points between contrasting regions involved in exchange or at the end points of natural transportation routes between such regions (Burghardt 1971; Hirth 1978). Additionally, they should form at critical nodes along such routes, such as bulk-breaking points or the juncture of different types of transport routes (Burghardt 1979; Bairoch 1988). Further, because of multiplier effects of trade on social evolution (Jacobs 1969, 2000), cities should concentrate in areas possessing the greatest positive productivity differential along a transport route, thus ensuring larger amounts of fungible surpluses usable for trade.

The central role of trade as a spur to processes of urban origins and growth is explained Jane Jacobs (1969, 2000), an iconoclastic urban expert. She argues that urbanism is a natural form of human organization once a threshold of population density and social complexity has been reached, and that social complexity and population density, in turn, are functions of economic differentiation. By implication, the question of urban origins thus devolves into the question of how economic differentiation is created initially
and the question of growth becomes that of how and at what rate differentiation expands. This is where trade comes in.

According to Jacobs, economic diversity is created in the first place as a result of positive feedback loops initiated by a settlement’s capacity to generate exports by combining some of its imports and/or preexisting resources with human labor and capital. This generates economic diversity at the same time that it makes it possible for more and different imports to be acquired, some of which can again be used to generate additional exports. This process creates co-developments in the form of an increasingly large, skilled, and diverse workforce (i.e., human capital), and this, in turn, creates the potential for further economic diversification by adding new types of work and new ways of working. Taken in the aggregate, social evolution is thus an open-ended and self-amplifying process, since each new economic differentiation constitutes the basis from which further diversity, both social and economic, can emerge. As both work and diversity expand, so does population density within the affected settlements. This increase commonly takes place at the expense of nearby rural populations, which is why developing cities are always the economic and physical shapers of their hinterlands.

Once founded, the key concept to understand how cities grow, in turn, is that of Circular and Cumulative Causation. This is an idea first articulated in the fifties by the economist Gunnar Myrdal (1957) and later elaborated, expanded, and formalized by Allen Pred (1966) for economics. The concept came to the attention of researchers interested in the rise of early civilizations soon thereafter through Jacobs’ seminal book, The Economy of Cities (1969). At its simplest, this involves the recognition that forces of production and urbanization are interlocked in a circular process whereby a change in one causes changes in the other which “… move the system in the same direction as the first change, but much further.” (Myrdal 1957: 13). As explained by Krugman (1995: 49) and Jacobs (1969, 2000), the most important mechanism whereby this takes place is import substitution.

After a regional economy grows beyond a critical point by means of the mechanisms described by Jacobs, it becomes profitable to replace imports of some commodities subject to scale economies with local production. This substitution will further expand urban employment, drawing in workers from the countryside and other regions; and in so doing will also further expand both the local market and the range of skills possessed by the population. The reasons for this have to do with the multiplier effects of increases in productive capacity. One is the creation of linked industries providing production inputs to the initial industry (backward linkages) or adding further value to semi-finished goods produced by those industries (forward linkages) (Pred 1966: 25–26). Another effect is the development of related work in sectors providing needed services, such as housing for workers, transportation, security, record keeping, etc.

In due course, the operation over time of these interrelated multipliers combine to create the enlarged population/market size necessary to induce further rounds of import substitution processes. As this process is repeated on an ever-larger scale, a circular (or, more precisely, spiral) relationship is created between population growth, market size, the range of productive activities that a region possesses, and the efficiency level of those activities: production is highest and most efficient where population and markets are larger but markets are bigger where production is greater, so that city-led regional growth (or decline) always takes the form of a self-reinforcing snowball or cascade effect (Krugman 1995: 49).
Early Mesopotamian Urbanism: Why?

What can we learn from these modern economic models of urban process that will help us better understand the forces at play at the onset of early Mesopotamian civilization? Two lessons come immediately to mind. First, trade and changes in commodity production and labor organization are as likely to have been fundamental agents of change in antiquity as they proved to be in modern times. Second, processes of circular and cumulative causation are also likely to have been as consequential in antiquity as they are today, so that if a region gains an initial advantage, those processes will concentrate new growth and its multiplier effects in the already expanding region rather than elsewhere (Malecki 1997: 49–50). Given this, what needs to be elucidated are the forces that set trade (and its multiplying ramifications) into motion in the first place.

In the case of early Mesopotamia, the trigger was provided by Cronon’s “natural environment,” referring, more specifically, to the combination of the enduring geographical framework of the Tigris–Euphrates alluvial delta and the unique environmental conditions that prevailed in the area during the fourth millennium BC. Geography was important because the absence from the southern Mesopotamian landscape of many materials necessary for the creation and maintenance of highly stratified social systems (most importantly roofing-grade timber, wood, base and exotic metals, various types of semiprecious minerals and exotic stones) made it probable that early southern elites would use trade as one of their most important tools to legitimize their unequal access to power and privilege. At the same time, geography provided those societies an enduring and irrevocable advantage over their neighbors in the form of lower transportation costs based on water transport.

Throughout their history, the cities of the Mesopotamian alluvium were situated, in effect, at the head of an enormous dendritic transportation system created by the north-to-south flowing Tigris and Euphrates rivers. This allowed them to procure information, labor, and commodities from areas within the vast Tigris–Euphrates watershed more efficiently than any potential upstream competitors or rivals away from the rivers. The crucial edge of southern cities lay in their ability to import needed commodities in bulk from faraway resource areas in the surrounding highlands at low cost, transported downstream on rivers by means of rafts or boats. Of equal importance, the network of canals surrounding Mesopotamian cities and connecting them with the main courses of the rivers allowed them to move bulky agricultural commodities across their immediate dependent hinterlands with great efficiency. In contrast, the land-locked societies of the Mesopotamian periphery had to rely wholly on less efficient modes of overland communication, such as donkeys or carts, both for their long-distance exchange needs and for the movement of subsistence resources across their immediate hinterlands.

Environment, in turn, gave southern societies equally important material advantages over surrounding areas. Historically, one such advantage was provided by the higher yields and reliability of the southern agricultural base. Modern agricultural data (summarized in Weiss 1986 and Wilkinson 1990) and inferences from ancient cuneiform documents (Jacobsen 1982, but see the reservations of Powell 1985) show that, under conditions of controlled irrigation, the alluvial landscape of southern Mesopotamia could be, on average, about twice or thrice as productive as the rain-fed agricultural regimes characteristic of neighboring societies. Additionally, the reliability of that production was
greater than that typical of dry-farmed areas at the periphery of Mesopotamia, which are subject to substantial and unpredictable spatial and temporal variations in rainfall (Perrin de Brichambaut and Wallen 1963: figs. 2–3; Erinç 1950; Turkes 1996).

These advantages would have been particularly pronounced at the time of initial urban emergence in southern Mesopotamia. A wide range of paleoenvironmental and geomorphological data (summarized in Hole 1994; Algaze 2001a: 202–03) suggests that the climate was wetter during the fifth millennium BC, and continued to be so through the first half of the fourth millennium. This meant that marginal areas of the Mesopotamian alluvium that are today unproductive because of insufficient water or lack of drainage would, at the time, likely have been integrated into fluvial networks draining into the sea (Hoelzmann et al. 1998:47). Additionally, parts of the alluvium that today receive no summer precipitation whatsoever would have been affected by summer monsoonal rains of Indian Ocean origin that today skip the northern edge of the Persian Gulf, but which had a more northerly track during the fifth and early fourth millennia (Sirocko et al. 1993; Petit-Maire, Sanlaville, and Yan 1995).

Lasting until the end of the Early Uruk period and, possibly, until the beginning of the succeeding Middle Uruk, summer precipitation must have had a very significant economic impact on the development of the earliest urban polities in the Mesopotamian alluvium. Joy McCorriston (personal communication 2001) suggests that among the most important of these would have been an expansion in the availability of animal forage at precisely the time of greatest need, particularly in the form of (C4) grasses (tropical, summer rainfall adapted): Panicum turgidum, Pennisetum divisum, Paspalum sp., Cymbopogon sp., Hyparrhenia sp., Heteropogon sp., etc. A second effect of summer precipitation would have been a substantial increase in the productivity of date palm grove cultivation.

These climatic advantages were compounded by advantages stemming from the nature of the Tigris–Euphrates fluvial system through the entire fourth millennium, which was very different from that typical for the historical periods. This realization has emerged only recently as a result of new work by Robert McC. Adams and Jennifer Pournelle (2003a, 2003b) using newly declassified satellite images (CORONA) to correlate the location of previously surveyed fourth millennium sites in southern Mesopotamia and relict watercourses visible in the images. They concluded that the two rivers formed a single dynamic network of anastomosing channels at the time of early urban emergence in the area, later separating to their historically known discrete courses after the fourth millennium. Thus, yield differences between alluvial Mesopotamia and its neighbors would have been greater during the Uruk period than was the case thereafter because waters of the two rivers were likely to commingle at flood stage in the northern parts of the alluvium, where their courses came closest, allowing for much larger areas than exist at present where various types of high-value vegetables and fruits could be produced in late spring and possibly early summer, by means of simple flood-recession irrigation (Sanlaville 1989:24).

Another natural advantage of southern Mesopotamia over neighboring areas was that the south possessed a greater variety and denser concentrations of exploitable resources found in complementary ecosystems. These resources are in fact often depicted in Uruk period iconography. They included (1) subsistence grain from the irrigable alluvial plain (e.g., Amiet 1961, Pl. 44: 639), (2) fruits, vegetables, and flax (used for
textiles) from cultivated gardens and orchards near the rivers (e.g., Amiet 1961, Pl. 16: 266), (3) extensive pasture for sheep, goats, and cattle created by fallow and recently harvested grain fields (Amiet 1961, Pl. 41: 618), and (4) abundant fish (e.g., Amiet 1961, Pl. 13bis: g; Pl. 15: 260), fowl, wild animals (e.g., Amiet 1961, Pl. 40: 609), and reed products obtainable in freshwater marshes, brackish lagoons, and estuaries at the intersection between the Tigris–Euphrates river delta and the head of the Persian Gulf.

In fact, resources from biomass-rich marshes would have been both particularly accessible and plentiful during the fourth millennium – accessible, because mid-Holocene sea-level rises at this time brought the many resources of the gulf’s littoral zone into close proximity to growing Uruk population centers (Sanlaville 1989; Pournelle 2003a–b) and plentiful, because the increased rainfall in the Tigris–Euphrates headwaters typical for the first half of the fourth millennium would have significantly increased the extent of marsh and lagoon areas in southern Mesopotamia. In turn, that increase in area was compounded by an increase in the productivity of Mesopotamian littoral resources of the time. The more northerly track of the Indian Ocean summer monsoon, noted earlier, must have increased the upwelling of nutrient-rich sediments and water oxygenation within the extensive marshes and lagoons surrounding Uruk population centers (Reichart et al. 1997), thereby substantially enhancing the biomass density available within the fourth millennium southern Mesopotamian marsh ecosystem. Additionally, because the leaves of young reed shoots can be used as fodder for animals (Pournelle 2003b), the margins of the greatly enlarged mid-Holocene marshes of southern Mesopotamia would have supported particularly large herds of cows, oxen, and sheep, providing both a uniquely ample supply of protein-rich dairy products (Englund 1995), and an important activity generating economic and occupational differentiation. The close connection between cattle and marsh resources postulated here for the Uruk period is, in fact, amply documented in the iconography of the time, which repeatedly depicts the sorts of bundled reed huts typical for marsh edges in southern Mesopotamia used as cow barns (e.g., Amiet 1961, Pl. 42: 629, 632 [seals] and Amiet 1961, Pl. 42: 623 [gypsum trough]).

In contrast, the high plains and intermontane valleys on the periphery of the Mesopotamian lowlands in the fourth millennium were unaffected by the monsoonal rains of the time and offered no comparably varied, resilient, or dense concentrations of subsistence resources beyond a once-a-year crop of dry-farmed grain and ample pasture lands which, though suitable for sheep and goats, could not have supported comparably dense bovine herds.

The consequences of the geographical and environmental advantages just noted are clear. On the one hand, advantages in productivity and resilience of their environmental framework meant that Uruk elites could extract larger surpluses per unit of labor than their counterparts elsewhere and they could do so with greater reliability and predictability. In addition, inherent advantages of water transport meant that Uruk elites could mobilize surpluses from their immediately dependent hinterlands at lower cost than their competitors, that the extent of those dependent hinterlands would naturally be larger than those of land-locked competitors, and that they could procure resources and information from a much vaster area at much lower cost than their contemporaries. Taken together, these advantages gave emerging elites in Uruk polities a substantial lead over their peers in neighboring areas in the amounts and variety of information and resources at their disposal, the size of the labor force at their command, and the productivity of those
laborers. Under these conditions, as economic geographers remind us, trade is the logical outcome.

**Early Mesopotamian Urbanism: How?**

The concept of circular and cumulative causation, discussed earlier, allows us to visualize a still highly speculative, though ultimately testable, scenario to account for the precocious urban takeoff of southern Mesopotamian societies in the fourth millennium. Following the insights of Jacobs, this scenario focuses on how economic differentiation could have been created in the south. For heuristic purposes, this hypothetical evolving process is divided here into a number of discrete stages, although substantial overlaps must be presumed to have existed between them.

The initial stage of the growth of southern economies would have taken place during the late fifth and early fourth millennia – a time when the geographical and environmental framework of southern Mesopotamia created a mosaic of very different but easily exploited resource endowments across what is today the Mesopotamian alluvium. In its northern portions, gravity-flow irrigation and increased water tables would have made grain cultivation and horticulture more profitable, whereas areas nearer the gulf were better situated to exploit its biomass-rich marshes, lagoons, and estuaries. Inadvertently, this setting provided the initial impetus for burgeoning trade between polities exploiting these varied economic resources. Each of these polities would have naturally specialized in the production of a small number of crops or commodities for which it had a comparative advantage, owing to its location within the alluvial ecosystem. Products traded in this initial stage would have included (1) woven and dyed textiles, goat-hair products, leather goods, dairy fats, and other pastoral resources distributed by polities situated at the margins of the better-watered parts of the alluvium, where they would have enjoyed preferential access to pastoral and nomadic groups producing these various commodities, (2) flax-based textiles, garden crops, and grain produced by polities in the northern portion of the Mesopotamian alluvial plain, where the combined flow of the Tigris and the Euphrates made irrigation agriculture and horticulture both more likely and more profitable, and (3) dried, salted, and smoked fish, various types of fowl, reeds, and other marsh or littoral resources preferentially produced by polities near the Persian Gulf littoral.

A second stage in the process may have started already by the Middle Uruk period and would have been marked by an emerging elite awareness of the social implications of the trade patterns in place until that point in time. In this stage, processes of competitive emulation would expedite the diffusion of technologies and practices that were initially developed by individual centers exploiting specialized niches but soon came to be perceived as highly advantageous by many of the centers in competition. This naturally led to a decrease in regional specialization within the alluvium as each competing polity used the material surpluses and human skills acquired during the earlier stage to replace some imports from nearby centers, or possibly even from foreign areas, by creating their own productive capacities for those products, thus setting in motion the further growth spurt that accrues from the import-substitution mechanism discussed earlier.

The third stage of the process is datable to the Middle and Late Uruk periods and would have been characterized by heightened competition between alluvial polities that had by now achieved broadly comparable productive capabilities. Since such polities no
longer had much to offer each other in terms of exchange, this stage was characterized by a significant expansion of external trade between individual alluvial cities and neighboring areas. Accordingly, ongoing import substitution processes in the south would have now focused largely on the replacement of foreign commodities. Enhanced foreign trade at this point was made possible by the domestication of the donkey and its widespread adoption as the main mode of overland communications across southwest Asia (Wright 2001: 127), which allowed southern societies for the first time to export in bulk locally produced goods.

As external trade became increasingly important in the Middle and Late Uruk periods, various types of southern outposts were established at strategic locations of significance for transport across the Mesopotamian periphery, principally but not solely, at the intersection of the north-to-south flowing rivers and the principal east-west overland routes across the high plains of northern Mesopotamia (Algaze 1993, 2001b, 2005; but see Rothman [ed.] 2001 for a compilation of views about the nature and function of the outposts). While these outposts may have served in part as outlets for displaced population from the south (Johnson 1987/88, Pollock 1999; Wright 2001; Schwartz 2001), their carefully selected locations suggest that they also served as collection and transshipment points for the increasing amounts of peripheral commodities imported into the alluvium in the later part of the Uruk period, and as distribution points for alluvial exports (Algaze 1993, 2001a, 2005).

It is easy to visualize the role that the still partly hypothetical patterns of trade just described would have had in the emergence of Sumerian civilization if we focus our attention on the long-term multiplier effects of the associated import substitution processes. These can be easily documented in the archaeological record of Late Ubaid and Uruk period Mesopotamian societies. Perhaps one of the earliest examples in the record is the partial substitution of imported flint for locally manufactured clay sickles, a process that starts already in the Late Ubaid period and continues through the various phases of the Uruk period (Benco 1992). Another example is provided by metals, first attested in the south by the end of the Ubaid period (Moorey 1994: 221, 255–58). Initially, metal goods must have been brought into southern Mesopotamia as fully finished products imported from metal-producing highland regions of Iran and Anatolia where metallurgical technologies were first developed (Kohl 1987: 16; Stein 1990). By the Middle–Late Uruk periods, however, Uruk societies had already created their own metal-processing industries that relied instead on imports of only lightly processed ores and of semi-processed ingots of smelted copper as opposed to fully processed tools, artifacts, and objects of personal adornment.

Evidence of the shift from metal consumers to value-added producers in the south (still using, of course, partially processed imported resources) is provided by ores recovered at Warka and ingots recovered at Jebel Aruda, an Uruk colonial enclave along the Euphrates in Syria, as well as by metal-processing installations identified in Uruk sites both in and out of the alluvium (Algaze 2001a: 208–209 for specific references). By the final phase of the Uruk period we also get textual corroboration for the shift. Metal objects figure prominently in many of the earliest economic texts of the time as well as in the much rarer, but critically important, contemporary lexical lists (Uruk IV Script; see Nissen 1986; Englund 1998). Additionally, the pictogram for a smith (showing a smelting
furnace with attached blowpipes) is also attested in the earliest Archaic Texts (Moorey 1994:243).

By far the most important case of import substitution processes in the south is provided by the adoption sometime by the Late Ubaid period of wool-bearing breeds of sheep initially developed in the highlands surrounding Mesopotamia (Davis 1984; Sherratt 1997: 539). Because such sheep are not indigenous to the lowlands, wool must have been initially introduced into the south as an import from the surrounding highlands. But wool and woolly sheep did not remain imports for long. As Joy McCorriston (1997) has recently noted, archaeobotanical, zooarchaeological, and textual data from various Uruk period sites shows that by the second half of the fourth millennium these imported commodities had been thoroughly integrated into the southern economy. This took the form of a fast-growing indigenous textile industry based on woven woolen textiles capable of being dyed that for all practical purposes replaced the less efficient and less colorful flax-based textiles that had constituted the bulk of local production in the south until then.

In spite of their late start, southern producers of woolen textiles soon surpassed their highland predecessors and competitors in both scale and efficiency. Several factors help account for this. The first is that by integrating the sheep into the agricultural cycle of grain, the south possessed as much fodder as the highlands, so that no dietary disadvantages accrued to the sheep as a result of their introduction to their new man-made habitat. The second is that the south had comparative advantage in access to pertinent natural dyes. This is another point recently raised by McCorriston (1999, 2001: 222 and pers. comm., 2001), who notes that many of the dyes used in conjunction with wool in the area in antiquity could be derived from desert or garden plants available in or around southern Mesopotamia, such as *Chrozophora tinctoria*, *Arnebia tinctoria*, *Papaver* sp., *Crocus* sp., *Salicornia* sp./*Cornulaca* sp. and *Punica granatum* (pomegranate), from plants available in the high plains of Syro–Mesopotamia and easily accessible by Uruk colonists in the area, such as safflower, or from products that could only be obtained from the Persian Gulf or through Gulf-related trade routes, such as various types of marine gastropods and indigo. In contrast, the highlands surrounding southern Mesopotamia were devoid of most plants from which usable dyes could be extracted, save for walnut. The third factor is that, for reasons already explained, southern societies possessed larger pools of labor available for textile work. From the beginning, these larger pools of workers appear to have been organized in ways that allowed for greater efficiency and superior craftsmanship in the production of textiles. This is inferable from contemporary depictions of presumably female workers (pony-tailed figures) attending horizontal looms (Amiet 1972: nos. 673–74; 1980: nos. 319–20). These images suggest that from its very beginnings, the woolen textile industry of the southern lowlands was based on state-organized establishments staffed by dependent women, such as we know existed in most Sumerian and Babylonian cities during the third and early second millennia BC (Jacobsen 1970 [1953], Maekawa 1980; Waetzoldt 1972).

Be that as it may, the shift from linen to wool as the primary material for textile manufacture in the south and the closely-related development of state-sponsored weaving establishments during the fourth millennium present us with a textbook case illustrating the many multiplier effects that commonly attend the introduction of new industries and increases in productive capacity, noted earlier. Particularly noteworthy would have been the forward and backward linkages related to the start of industrial-scale weaving in the
south. Examples of the former are provided by the fulling of semi-finished woven textiles with oils and alkali and the dyeing of fulled cloth. Both of these practices are well attested in later third millennium textual documentation and both require a substantial input of value-adding labor and new resources (Potts 1997:95; McCorriston 2001:222). Examples of backward linkages, in turn, are provided by a variety of labor-intensive activities that contributed necessary inputs to the weaving establishments but largely took place away from them. Minimally, these included pasturing the sheep, washing, plucking and/or shearing, combing, and spinning the wool, separating the wool by quality, and delivering it to the various locations where state-organized weavers labored. No less important would have been other multiplier effects of the shift to industrial-scale textile production. Under Mesopotamian conditions, this would have required scores of bureaucrats to record, store, and redistribute the output, and also to supervise the housing of laborers and the distribution of subsistence rations.

The Evidence for Trade

The foregoing discussions follow the lead of economic geographers interested in the dynamics of urban growth and identify trade, both internal and external, as the engine of early Mesopotamian urban growth. There is, however, substantial disagreement about the importance of trade in general and long distance trade in particular to the processes of urban and state formation in southern Mesopotamia.

Many scholars reviewing data for southern Mesopotamian economies of the fourth millennium properly highlight the importance of local tribute extraction and intra-regional distribution of resources as key elements in that economy but either minimize the overall importance of long-distance trade to the socio-economic processes at work at the time (e.g., Frangipane 2001; Pollock 1999; Schwartz 2001:256; Weiss 1989) or presume that rises in long-distance exchange were a consequence of urbanism rather than a cause (e.g., Wright 1981a, 2001). Such views are flawed on two accounts. First, they fail to acknowledge the evidence for valuable imports in both the textual and archaeological record of Uruk sites. The case of metals and precious stones is particularly instructive. The earliest Archaic Texts, for instance, already include numerous references, noted earlier, to metals, which must have been imported into the south either as finished products or, most likely at this point, as ores and ingots. Similarly, some of the Middle and Late Uruk period southern colonial sites on the Euphrates yielded ample evidence for the import and in-situ processing of silver and copper ores and various types of exotic stones, including lapis (for specific references, see Algaze 2001b: 208–209), all presumably for re-export to larger Uruk centers in the south. Those centers, in turn, have also yielded some evidence for this wealth. Though examples abound, nowhere is this clearer than in the so-called Riemchengebäude structure found in the Eanna Precinct at Warka (Late Uruk: Eanna IV), which was literally brimming with many categories of imported exotic materials (for an inventory, see Forest 1999: 67–73).

Second, views that deny or minimize the early importance of trade consistently overestimate how representative the archaeological record really is. In fact, many of the articles traded in the Uruk period would have left few traces in the archaeological record. A case in point is the likely principal alluvial export of the time: elaborately crafted textiles. The same is the case, albeit for entirely different reasons, for many of the imports
into the alluvium at the time. No doubt, the most important (by volume) would have been timber and wood, which again tend to disappear from the archaeological record and can only be inferred indirectly from considerations of architectural needs (e.g., Margueron 1992). Other contemporary imports would have consisted largely of a variety of exotics and semi-precious commodities, principally metals, which are also unlikely to be preserved in representative amounts in the archaeological record of complex central sites, absent a destruction level.

The reasons for this are explained by Andrew Sherratt (2004), who notes that in truly complex societies sumptuary goods and metals will be distributed more widely across social hierarchies than in simpler societies, as such commodities become a medium of exchange capable of being converted into a wide range of goods and services. This naturally increases the likelihood that such commodities will be kept longer in circulation, that they will be transformed more often, when practicable (e.g., metals by melting), and that they will be passed on across generations more consistently. In situations where commodities circulate across wide social networks, excavations at single central sites, or, worse still, at the core of such sites, are increasingly unlikely to produce a representative sample of the scale and type of sumptuary commodities in circulation any one time. Regrettably, such excavations presently provide the bulk of the available evidence for the Uruk period in the alluvium.

**Issues of Scale: South versus North**

Where trade flows, its ramifications in the form of increasing social complexity and urbanism soon follow. Thus, the precocious development of southern Mesopotamia throughout the fourth millennium BC comes as no surprise. How unique development in the south was at this point becomes clear when we compare available survey and excavation data for both the nature of sites and of patterns of settlement in the alluvium against comparable data from neighboring regions, and particularly from Upper Mesopotamia.

For the south, available survey data (Adams 1981; Adams and Nissen 1972; Wright 1981b; for a reworking of the data, see now Kouchoukos 1998: 230–249; Pollock 2001) reveal that both absolute population levels and relative agglomeration rates were significantly higher throughout the various phases of the Uruk period than anything that existed in any one coherent area of likely cultural interaction of the Mesopotamian periphery (Kouchoukos 1998:Tables 5.4–5.6, Fig. 5.9). In fact, surveys document multiple interacting urban sites within the surveyed portions of the alluvium throughout every phase of the Uruk period, all situated alongside canals and within relatively short distances of each other and each positioned at the apex of a variegated settlement structure.

---

1 Note, however, that Wilkinson’s detailed and particularly systematic survey of the Upper Jezira plains west of the Tigris in northern Iraq show what appear to be higher overall regional population densities in that area than in the south (Wilkinson 2000: Fig. 5). It is unclear whether this represents a real pattern or whether it is a consequence of depressed site counts in the south due to sedimentation (Wilkinson 2000: 244). Although I am inclined to the latter position, it is certain that in either case the south still had a much greater proportion of its overall population living in agglomerated settlements and that these settlements were situated at much shorter distances from each other than northern centers.
This pattern was in place already by the Early Uruk period, when at least four sites across the alluvium were 40 ha in extent or larger (Uruk, Eridu, Site 1237, and Site 1306) and multiple other sites were in the 20–40 ha range (Adams 1981; Wright 1981b). Development in the southern alluvial lowlands reaches its peak by the final phase of the Uruk period, when the site of Warka reaches the extraordinary size of 250 ha (Finkbeiner 1991). Warka was surrounded at this time by numerous dependent villages and towns, totaling a minimum of 280 ha of further occupation (Adams and Nissen 1972). Again, various sites in the 20–40 ha (Nippur, Site 1172, Site 125) and 40+ ha (Site 1306) range existed elsewhere across the alluvium at this point (Adams 1981).

The sites just discussed are likely to be only the tip, so to say, of the Uruk period settlement iceberg in southern Mesopotamia. Two reasons account for this. First is the fact, already noted, that geomorphological processes operating in the southern alluvial plains of Tigris and Euphrates rivers are particularly likely to obscure a large proportion of the smaller Uruk period sites in the south. Additionally, in places, particularly dense Late Holocene alluvial deposits may even obscure some of the larger Uruk sites in the south. This possibility is raised by J. Pournelle (2003a, 2003b), who notes that, in some cases, satellite imagery shows that multiple small nearby Uruk period sites recorded by Adams as independent settlements were instead parts of larger, single shallow settlements partly covered by alluvial deposits.

Second, and more concretely, a number of sites exist outside of the areas surveyed by Adams and others that were occupied during one or more phases of the Uruk period. These sites are not considered in recent reviews of the nature of Uruk period settlement in southern Mesopotamia (e.g., Algaze 2001a; Pollock 2001; Wilkinson 2000) but several are likely to have been quite substantial at the time. Foremost among these are Umma and the nearby site of Umm al-Aqarib.2 During a recent brief visit to Umma, McGuire Gibson (personal communication, 2001) reports Uruk pottery spread widely over the surface of the site. More tellingly, numerous Archaic Texts recently plundered from either (or both) of those sites appear immediately comparable to the earliest examples from Warka (R. Englund, personal communication, 2001). At a minimum, these tablets attest to the economic importance of the Umma area in the Late Uruk period. However, since at Uruk these tablets are part of a wider urban assemblage of great extent and complexity, their presence in the Umma area argues for a similar context.

Though circumstantial, this evidence suggests that Umma and its satellites may have been second only to Warka itself in terms of urban and social development in the Late Uruk period. Buttressing this possibility is a glaring anomaly in the settlement data for Late Uruk southern Mesopotamia as presently known: the largest site (Warka) is four times as large (i.e., populous) as second-tier settlements (e.g., Site WS 1306). This is anomalous because analyses of modern urban systems show that urban populations arrange themselves in rank order by size in predictable ways (“Zipf’s Law”), with each tier of settlement being anchored by a site roughly double in size that of the largest settlement of the preceding tier (Krugman 1996b). If comparable rank-size behavior characterized the ancient Mesopotamian urban world, as I would expect to be the case, then 60-hectare range sites such as WS 1306 should not represent a second Uruk Period settlement tier, but

---

2 Both Umma (WS 197) and Aqarib (WS 198) were at the edge of Adams’ 1968 survey area, but could not be properly surveyed at that time because of extensive sand dunes covering the area (Adams and Nissen 1972: 227-28). The dunes have since cleared the area.
rather a third. The missing (second) tier should be anchored by a site roughly half the extent (population) of Warka. I expect that Umma will eventually be found to represent this missing tier, and that further work at that site will eventually show it to have been somewhere in the range of 120 ha in the Late Uruk period.3

The long sequence of urban growth in the southern Mesopotamian alluvium throughout all phases of the Uruk period contrasts starkly with the overall developmental trajectory of contemporary northern Mesopotamian societies. To be sure, as Henry Wright (2001:145) presciently noted, both sequences similarly start the fourth millennium with an initial burst of settlement growth and expansion of social complexity which, in the north, lasts until the end of the Middle Uruk period, ca. 3400 BC or so. This has become evident only recently in Upper Mesopotamia as a result of new excavations at Tell Brak (Oates and Oates 1997; Emberling et al. 1999; Emberling and McDonald 2001) along the Jagh Jagh branch of the Upper Khabur River in Syria, new excavations at Nineveh along the Upper Tigris river in Iraq (Stronach 1994), new surveys at Tell el-Hawa and its environs (Wilkinson and Tucker 1995), in the Jebel Sinjar Plains of northern Iraq, and older surveys of Samsat and its environs (summarized in Algaze 1999), in the Upper Euphrates area of southeastern Turkey.

This new body of work shows that the scale of individual sites in disparate areas of the northern Mesopotamian plains during the first half of the fourth millennium was roughly comparable to that of contemporary sites in the southern Mesopotamian alluvium. Tell Brak, for instance, grew to a minimum of 65 ha ( Emberling et al. 1999) in the so-called “northern Middle Uruk” period, and may have been even larger depending on whether or not the intervening area between the main site and nearby contemporary suburbs was occupied (Oates 2001, H. Wright, personal communication 2004, J. Ur, personal communication, 2003). Brak was thus broadly similar in extent, and possibly even larger, than Uruk itself and Site 1306 (Adams’s “Early/Middle Uruk” phase; Adams 1981). Nineveh too is likely to have been substantial at this time. Its most recent excavator, David Stronach (1994) gives a preliminary estimate in the 40-ha range. Hawa is reported to have been in the 30+ ha range at this time (Wilkinson and Tucker 1995), and Samsat and Hamoukar were about half that size (Algaze 1999; Ur 2002).

Similarities in the scale of individual sites in northern and southern Mesopotamia, however, mask important differences in the complexity of the settlement systems of both areas as a whole. Even at their peak in the “northern Middle Uruk” period, Late Chalcolithic societies of the north hardly equaled their southern counterparts in complexity. This is inferable from comparisons of available survey data bearing on the density and hierarchical structure of settlement grids surrounding large settlements in both areas throughout the fourth millennium. Pending the publication of recent surveys around Brak conducted by H. Wright and his colleagues, the best data we have for the north is

3 Further candidates for significant Uruk period sites or site complexes outside of the present boundaries of systematic survey in the south include Girsu, where French excavations in the late 19th and early 20th centuries uncovered substantial Uruk period remains (Parrot 1948), and the nearby but smaller site of Surghul, where Uruk deposits also exist (E. Carter, personal communication 2001). Since both of these sites were central settlements in the later third millennium kingdom of Lagash, it is not entirely farfetched to see them as parts of a smaller earlier polity of the Uruk period.
derived from systematic surveys for the Hawa and Samsat environs by Tony Wilkinson, which show that during the first half of the fourth millennium both sites were surrounded by a corona of uniformly small village or hamlet-sized sites (Wilkinson 1990, Algaze 1999; Wilkinson and Tucker 1995:fig. 35, top). This compares unfavorably with the more complex settlement grids of variously sized dependent settlements that surrounded contemporary (Early/Middle Uruk) urban centers in the south (Adams 1981; Pollock 2001). Further, surveys of the Hawa and Samsat environs show that a more complex three-tiered settlement pattern structure appears in their vicinity only after the onset of contacts with the Uruk world, not before (Algaze 1999; Wilkinson and Tucker 1995:fig. 35, bottom).

More important still is a further difference between northern and southern Mesopotamia. Large Late Chalcolithic settlements in the northern plains such as Nineveh, Hawa, Hamoukar, Brak, and Samsat were situated in different drainages and were separated from each other by hundreds of kilometers. Thus, they were largely isolated from one another in terms of day-to-day contacts. This was not the case in the south, where multiple competing settlements connected by waterways existed within short distances and easy communication (via water) of each other.

In light of the above, it should not be surprising to find sharp differences in the overall developmental trajectories of both areas through the fourth millennium. Most salient among these is that in the north, unlike the south, the initial burst of growth and development was not sustained for long. Data from Nineveh, Hawa and Samsat are unreliable on this point, but new excavations at Brak show that the settlement contracted in the second half of the fourth millennium (Emberling et al. 1999: 25-6; Emberling 2002), just as the expansion of southern sites such as Warka reached their Late Uruk peak. Available data are not precise enough to discern whether the contraction of proto-urban fourth millennium Brak was caused by the intrusion of Late Uruk elements into the site, as Emberling (2002) now suggests, or whether the intrusion took advantage of a preexisting process of decline. Either way, the contraction of Brak in the final quarter of the fourth millennium meant that urban centers in the alluvium of Late Uruk date were significantly larger than contemporary Late Chalcolithic polities in the Mesopotamian periphery. In fact, at 250 ha, Late Uruk Warka is likely to have been many times larger than any contemporary peripheral competitor. The fact that this huge differential developed at precisely the time of the maximum expansion of the Uruk colonial network is unlikely to be a mere coincidence.

Developments at Brak during the second half of the fourth millennium (Late Uruk) appear representative of northern Mesopotamia as a whole, and clearly indicate the start of a centrifugal process culminating in the widespread ruralization of the northern plains by the end of the fourth and the transition to the third millennium. This is suggested by the results of surveys of portions of the northern plains within modern day Iraq, Syria, and southeastern Turkey, which consistently show that sites across the area dating to the fourth-third millennium transition (“Kurban V” and “EBI” along the Euphrates and “Painted Ninevite V” along the Khabur and Tigris basins) were uniformly small villages or hamlets (data summarized in Algaze 1999: Table 3). By this time, the few indigenous centers that had existed in the preceding period had vanished, and such centers would not reappear until the final phases of the Ninevite V period, sometime in the second quarter of the third millennium (Schwartz 1994; Weiss 1990; Wilkinson 1994).
In contrast to the aborted proto-urban experiment of the north, urbanism in the south continued to flourish and expand not only though the Late Uruk period but throughout the fourth/third millennium transition (Jemdet Nasr/Early Dynastic I) as well (Adams 1981, Postgate 1986). In fact, by the first quarter of the third millennium, at a time when no urban centers are known to exist in the north, the urban spiral of the south continued unabated (Early Dynastic I): older sites such as Ur, Kish Nippur, Abu Salabikh, Warka, and, possibly, Umma grew further, and new cities were founded across the alluvium, including most notably Lagash (Al-Hiba) and Shuruppak (Fara) (Adams 1981; Wright 1981b; Gibson 1972). Warka reached 600 ha in extent at this point (Finkbeiner 1991) but this was no longer exceptional; Al-Hiba situated at the edge of the easternmost marshes in the alluvium was almost as large (Carter 1985).

The Synergies of Civilization

Multiple consequences would have arisen from the just-discussed differences in population density and distance between polities typical of southern Mesopotamia and areas on its periphery throughout the second half of the fourth millennium. These consequences represent in effect socio-evolutionary synergies that help explain why the earliest urban and state-level societies of southwestern Asia appeared in southern Mesopotamia and not elsewhere.

The first synergy arises from the greater concentration of polities that existed in the Mesopotamian alluvium throughout the 700 year or so duration of the Uruk period, as compared to neighboring areas. As Colin Renfrew and his colleagues (Renfrew and Cherry 1986) have repeatedly argued, the long-term presence of multiple polities within relatively short distances of each other invariably engenders important processes of competition, exchange, emulation, and technological innovation – processes that, as argued earlier, are archaeologically visible in changes in how commodities were produced in Middle and Late Uruk period Mesopotamia. The impact of these mutually-reinforcing processes has been explained by Robert Wright (2000:165-68), who notes that in situations where antagonistic but mutually communicative polities exist, social and economic innovations that prove maladaptive in any one society are likely to be weeded out more quickly than in less competitive settings. Conversely, innovations that prove advantageous are more likely to spread quickly across the various polities in competition, thus accelerating the pace of change of the system as a whole.

The second synergy arises from the greater proportion of the population of southern Mesopotamia that lived in towns and cities and their immediate dependent hinterlands through the Uruk period, as compared to the more dispersed settlement typical for neighboring areas at the time. This had many consequences. First, as Adam Smith (1954 [1776]: bk I, chaps. 1-3) noted more than 200 years ago, the assemblage of a critical mass of both producers and consumers is a necessary precondition for the division of labor and resulting economies of scale (below). Additionally, proximity between workers and employers lowers training costs and increases labor flexibility (Malecki 1997: 49), thus providing southern institutions quicker access than their competitors to skilled workers/builders/soldiers in times of growth and need. Second, increasing population density in towns and cities would have compounded the natural transport advantages of the alluvial environment by further efficiencies in transportation and communication arising from the increasingly compact arrangement of the inhabitants of the area throughout the
fourth millennium. One such compounding efficiency, which falls squarely in the realm of Cronon’s created landscape, was provided by the start of construction of minor artificial canals across portions of the southern alluvium through the Uruk period. These canals are partly visible in available CORONA imagery of the southern Mesopotamian alluvium, which only became available well after the publication of the final report of Adams’s surveys in 1981 (Pournelle 2003a, figs. 2, 8). The construction of such canals was probably highly opportunistic, and they often take off from natural avulsive nodes along the main river channels (Pournelle 2003b: 197, fig. 80). In addition to their agricultural role, these small canals extended the natural transportational advantages of the Mesopotamian landscape to areas beyond the natural flow of the rivers. In so doing, they also served to reinforce ongoing urbanization processes in the alluvium. This effect, no doubt inadvertent, may be inferred from studies that clearly link reductions in transport costs of agricultural commodities in traditional societies to the expansion of existing agricultural boundaries and the movement of population into cities (Fujita and Krugman 1995: 520). A final compounding efficiency has already been mentioned, and was provided by improvements at this time in the facility for overland movement in and out of the alluvium in the Middle-Late Uruk periods as a result of the introduction of domesticated donkeys and wheeled carts (Bakker et al. 1999). While these technologies were shared by a wide cross-section of contemporary ancient Near Eastern societies (Stein 1990; Kohl 2001), they must have impacted Uruk societies with particular intensity. Only in the south did advances in the efficiency of overland travel complement preexisting advantages derived from water transport, and only in the south did both of these, in turn, reinforce and facilitate advances in commodity production using task-specialized labor and in the ability to transmit information accurately across time and space (below).

The third synergy is closely related to the preceding and again arises from increasing density and compactness. This inevitably led to a multiplication of interactions between individuals in the increasingly urbanized landscape of southern Mesopotamia through the Uruk period. As interpersonal interactions multiply, information flow is enhanced. In turn, this radically enhanced the possibility that technological improvements and inventions would have taken place in the south at this time. Why this should be so is explained by Gerhard Lenski (1979: 16), a sociologist, and Joel Mokyr (1996: 71), an economist, who note that technological innovation is essentially a process of recombining existing elements of information, so that the rate of innovation is bound to rise as both the store and flow of information increase. At a minimum, this increase would have been exponential, because with each doubling of the number of people in contact within an urban grid, the number of possible vectors of interaction is squared. However, because each person in the grid would have actually possessed multiple elements of information capable of recombination at different times, and because interaction could have taken place between multiple individuals at any one time, the actual increase in the probability of innovation in the Uruk world would have more closely conformed to the parameters of a quadratic growth curve rather than to those of an exponential one (i.e., the number of individual pairs of elements of information potentially free to interact at any one time would rise foursquare with the squaring of their basic number).4

4 I am grateful to Douglas White (UC Irvine) for helping me understand the principles underlying power-law growth patterns and their applicability to processes of urban takeoff.
Technologies of the Intellect

As the web of interpersonal communications became increasingly dense in southern cities that, by the second half of the fourth millennium, were growing many times larger than neighboring population centers, the likelihood that advantageous inventions and innovations would arise and be quickly diffused was greatly enhanced. In the Mesopotamian case, this found expression in a variety of revolutionary technologies of social control that are as much part of the “created landscape” as the new irrigation canals. These new technologies largely fall in the realm of what the eminent social anthropologist Jack Goody (2000) has termed “technologies of the intellect.”

Perhaps the most salient of these ideational technologies was the systematic use of various types of dependent laborers receiving rations for the production of subsistence and sumptuary commodities and for building and agricultural activities. This amounts, in effect, to a “domestication” of human labor (Algaze 2001b), and represents a veritable turning point in the evolution of socially stratified societies across the world. Evidence for this revolutionary development in southern Mesopotamia is provided by (1) millions of Beveled Rim Bowls, which presumably served as ration containers, found at a variety of Uruk sites (Nissen 1988), (2) hundreds of attestations of signs for various categories of dependent laborers in the known corpus of Archaic Texts from Warka, and (3) at least one tablet that summarizes food rations given to different groups of male and female captives (Englund 1998:70, 178-79, fig. 66). This means that Uruk elites had more laborers at their command than competing elites elsewhere, that they could extract more energy from those laborers, and that they were better able to move them around as needed at little cost – an ability often identified as a key factor in economic development (Krugman 1995:19). More importantly, it also means that Uruk elites could organize laborers in nontraditional ways so as to take advantage of increases in productivity and other economies of scale arising from the specialized production of commodities. The clearest material evidence for this organizational quantum leap is provided by the well documented shift to standardized, mass-manufactured ceramics throughout the Uruk period (Nissen 1976) but comparable changes can be seen in the way other commodities were produced or procured at the time, for instance, wool (Green 1980; Nissen 1986) and metals (Nissen 2000). These patterns leave no doubt that economies of scale based on task specialization were being introduced in a variety of Uruk period productive activities.

A second ideational technology appearing at this time in the south consisted of new forms of information processing and record keeping that were more capable of conveying information across time and space than the simpler reckoning systems used by contemporary societies elsewhere (Algaze 2001a: 212-213). This process started in earnest in the Middle Uruk period with the introduction of complex reckoning devices such as impressed hollow balls filled with tokens capable on conveying information by combining numbers and images, and culminated in the Late Uruk period with the development in quick succession of, first, seal impressed numerical notation tablets, and, second, tablets bearing fully developed pictographic writing (Archaic Texts: Uruk IV Script) (Nissen, Damerow, and Englund 1993). The overwhelming majority of these latter tablets were simple accounting texts recording inflows and outflows of commodities. In essence, these tablets served the same function as the much earlier seal impressions on clay that had accompanied the movement of commodities across Mesopotamia since
prehistoric times. What is profoundly revolutionary about them, however, is that such commodity flows were being recorded in a manner that allowed for the expression of nuances of time, location, persons involved, and action effected, and that was transmissible through space and time. As the Assyriologist Robert Englund (1998, 2004) has noted, even the earliest pictographic tablets show that southern scribes had the ability to abstract and summarize detailed data about collections and disbursements of goods and labor in a form usable by themselves at a later time, by higher-level supervisory officials at any time, and by later generations of similarly trained bureaucrats (see also Hudson 2004 and Steinkeller 2004). In so doing, these tablets allowed early Mesopotamian decision makers, and the urban institutions they worked for, to deploy available labor and resources so as to maximize their future revenues and power.

But the importance of writing goes much further than accounting for resources at hand and planning for future gain. A small but critically important proportion of the earliest Archaic Texts consists of thematically and conceptually arranged word lists (“lexical texts”) that, no doubt, served as scribal training exercises. These compilations provide unique insights about many aspects of the material, social, and ideological world of early Sumerian urban dwellers that are not generally referenced elsewhere (Englund 1998, 2004). More to the point, they presuppose the existence of a formally constituted scribal class dedicated to the transmission of knowledge across generations. Aided by scribes, early Sumerian elites and institutions would have had better and more detailed access to the accumulated knowledge of earlier generations than their rivals in neighboring areas, where the lack of comparably accurate and efficient forms of communication systems meant that the past would only be known through fallible human memories and ever mutable oral traditions (Goody 2000). Bluntly put, this meant that by the final phase of the Uruk period, the web of interpersonal communications across the Uruk World was being thickened by interaction not only between the living but also, and for the first time in human history, between the living and the dead. Equally important, because writing is a form of “cognitive scaffolding” or “external memory” that permits individuals to perform cognitive tasks above and beyond those normally possible by the unassisted brain (Mouck 2004), the presence of a scribal class in Uruk cities guaranteed that, as a group, Sumerian elites would have been more likely than their peers to possess the problem solving tools and institutional memory that are needed to efficiently integrate larger populations and more diverse territorial realms, to successfully react to recurring environmental perturbations and social threats, and to profitably take advantage of stochastic opportunities for gain.

In short, cumulative innovations in the way knowledge was gathered, processed, and transmitted in early Sumerian culture throughout much of the Uruk period provided the nascent urban polities of southern Mesopotamia at the time with what, arguably, was one of the most important competitive advantages that they possessed over contemporary polities elsewhere, in which comparable breakthroughs in accounting, accountability, classification, and access to current and past information appear to have been absent.

The Urban Revolution Revisited

The synergies and ideational innovations just discussed represent multiple and interrelated facets of a single phenomenon: advances in the efficiency and intensity of social
interactions possible within and between southern Mesopotamian societies of the fourth millennium above and beyond those practicable in neighboring areas at the time. These advances are key to understanding the Sumerian takeoff because, as the sociologist Amos Hawley (1986: 7) notes, historically, human settlements have exhibited a tendency to grow to the maximum size afforded by the technology for communication and transportation possessed by the population. Improvements in the ability to move materials, people, or information inevitably lead to increases in mean aggregate settlement size. This correlation has been known since the time of Adam Smith (1954 [1776]: bk I, chaps. 1-3), who observed that gains in the efficiency of transportation and communication always act as a spur for economic specialization and growth in human societies. The reasons for this are explained by Hawley (1986: 65-66), who notes that social units engaged in specialized functions are necessarily spread over space, which naturally decreases the efficiency of information flow and increases the cost of value-added production and services. Thus, increases in communication efficiency and reductions in mobility costs always result in gains in economic specialization and differentiation – processes that, as noted earlier, are central to the origins and growth of urban societies.

It is not difficult to see how these processes apply to the fourth millennium Mesopotamian case. The Sumerian takeoff correlates both with enhanced communication efficiency and accuracy in the form of new reckoning and writing systems and reductions in mobility costs as population became concentrated, production facilities consolidated, and production itself standardized.

Conclusions: The Mesopotamian Conjuncture

If we are to understand why the balance of urbanization and social complexity in the ancient Near East shifted decisively to the southern alluvial lowlands of Mesopotamia in the second half of the fourth millennium BC, we must delineate the sequence of mutually-reinforcing necessary and sufficient conditions that came together in the south at that time but were absent (in the aggregate) from neighboring contemporary social groups. Only by so doing can we begin to elucidate why the Sumerian takeoff took place at all, why it occurred when it did, and why comparable developments failed to materialize in Upper Mesopotamia, an area that only a few centuries before had appeared as poised for takeoff as the south.

Early on the stage was set by advantages in productivity, reliability, and ease of transport inherent to the “natural landscape” of southern Mesopotamia. Absent in the aggregate from neighboring regions, these advantages can be considered as the initial set of necessary conditions in the conjuncture. No doubt, the most important of these advantages was ease of transport. This becomes clear when we contrast the developmental sequences of southern and northern Mesopotamia in the fourth millennium and the locational circumstances of the principal settlements in each area at the time. Those of the south, as already noted, invariably lined the banks of waterways. So, for that matter, did the bulk of the known large Late Chalcolithic settlements across the north. Nineveh, Brak, and Samsat, for instance, are all situated along the principal navigable waterways crisscrossing the area. Each controls a historical fording place where the principal east-west overland routes across Upper Mesopotamia intersect the rivers (Algaze 1993, 2005). Paradoxically, however, water transport, the same factor that fostered interaction between
early centers in the closely intertwined fourth-millennium fluvial system of the south, limited interaction in the north, where the principal waterways were quite distant from each other and not suitable for the construction of artificial canals.

The intervening plains across the north also impeded both interaction and agglomeration compared to the south. Even after the introduction of donkeys and carts in the mid-fourth millennium, limitations inherent to overland travel in the area imposed enduring natural limits to population agglomeration away from the rivers (Wilkinson 1995). Whereas geography in the south both permitted and encouraged linearly arranged agglomerations based on boat and raft transport, the geography of the northern plains encouraged population dispersal instead so as to maximize the amount of territory under cultivation. Thus, initially at least, the northern plains naturally tended to foster smaller agglomerations than were possible in the south, and significantly more dispersed ones as well.

Under these circumstances, a critical mass of compact and closely interacting polities such as existed throughout the Uruk period in alluvial Mesopotamia failed to form across the north in the fourth millennium, making processes of intra-regional exchange, competition, and emulation in that area both less necessary and less likely than was the case in southern Mesopotamia at the time. At the same time, however, northern societies would also have had both less need and less ability than their southern counterparts to engage in bulk external trade with its many social ramifications. Less need because northern societies were generally situated in areas closer to the principal required bulk resources, such as timber, which could therefore be obtained locally without substantial organization. Less ability, because the means of transportation available to northern societies away from the rivers simply did not lend themselves to the cost-effective movement of anything other than low-bulk, high-value exotics. In contrast, the rivers provided a particularly efficient mode of channeling and distributing both trade in exotics and bulk trade for southern societies.

Accordingly, the initial proto-urban social systems of the north were not likely to expand significantly in size beyond a certain threshold, because of difficulties in transport. Nor were they likely to significantly enhance their productivity, because they lacked the critical mass of population to permit much specialization of labor or to encourage the development of new, more complex technologies of communication, such as proved fundamental for the Sumerian takeoff. This left an indelible mark on the historical development of the north because those types of social synergies were in fact precisely what was required in order for northern social systems to successfully circumvent the inherent constraints of their geographical framework. Indigenous city-states comparable (in complexity, if not always in scale) to those that had thrived in the south since the fourth millennium did emerge across the Upper Mesopotamian plains sometime just before the middle of the third millennium (Weiss 1990; Wilkinson 1994), 800 years or so after the Sumerian takeoff. However, it was only by adopting forms of social and economic organization and writing systems derived from southern models (Postgate 1988) that Upper Mesopotamian polities of the Early Bronze Age were able to marshal the organizational efficiencies needed to overcome the natural friction of overland travel across the area that had prevented their Late Chalcolithic predecessors from forming enduring regionally organized societies. Bluntly put, the initial experiment with early social complexity in Upper Mesopotamia ultimately failed because urbanism in the north
was only possible as a created landscape: it only became viable as a result of innovations in communication and labor control created elsewhere. In southern Mesopotamia, on the contrary, urbanism was a logical outgrowth of natural and socially-created synergies that compounded and reinforced each other from the very beginning.

In the end it turns out that Wittfogel (1957) was right, but for the wrong reasons. Rivers were indeed central to the development of early Mesopotamian civilization, as he argued, but not so much as a source of irrigation water but because of their role as conduits of transportation for subsistence commodities, building materials, necessary resources, and sumptuary goods. After all, in Mesopotamia as elsewhere along other river basins where pristine civilizations formed, cities emerged not at random along the courses of the rivers but rather in fertile areas downstream, where a minimal threshold of access to local agricultural resources was ensured and where, more importantly, transport costs were lowest and access to diverse resources within the river’s watershed was highest (Bairoch 1988:12). The importance of rivers to the emergence and growth of many urban societies is elegantly explained by Felipe Fernández-Armesto (2001:182), a historian, who notes that “… civilizations of scale can only be built with concentrated resources. Resources can be concentrated only by means of good communications. And for almost the whole of history, humankind has depended for long-range communications on waterways.”

And yet, natural advantages derived from geography and the environment do not explain in and of themselves the crystallization of early Mesopotamian civilization – or that of any other pristine civilization for that matter. In the final analysis, environmental and geographical factors are only permissive, not prescriptive. Whether individuals and groups react to environmental changes and take advantage of geographical possibilities, and how they do so, are always constrained by culturally determined perceptions of opportunities and threats at any one time. These, in turn, are partly shaped by available technologies and capital (both human and material). Moreover, the present is also shaped from the past by inherently unpredictable accidents and innovations that add an element of indeterminacy to any attempt at historical prognostication (or explanation). For these reasons, history displays a wide range of results of the interaction of societies and their environment, and this range can only become greater and more unpredictable as the density and intensity of social interactions grows in increasingly complex societies. Nonetheless, environment and geography do constitute important selective pressures that often impose an important measure of directionality on human affairs, as the historian E. W. Fox (1971, 1989) has repeatedly and persuasively warned us. The reason for this is explained by Joel Mokyr (1990), who notes that environmental factors commonly act as “focusing devices” limit the range of options that are perceived as viable by individual societies at any one time and that powerfully influence the direction that those societies take in their search for technological innovations.

Against this interplay between indeterminacy and directionality, the natural advantages of the southern Mesopotamian landscape merely provided a backdrop wherein some social responses became more likely than others. In light of the diversified but dispersed resources prevalent in southern Mesopotamia throughout the late fifth and fourth millennia BC, and given the naturally low mobility costs of the area, one of the most probable such responses was for pre- and proto-historic elite individuals and groups to specialize in the production of a limited number of commodities for which they had comparative advantage owning to their location within the alluvial environmental mosaic.
of the time and to engage in trade with differently specialized local rivals from relatively early on. By the same token, the absence of important necessary resources from the Mesopotamian environment, most notably roofing-grade timber and metals, also made it likely that early southern elites would seek to engage in trade with foreign counterparts in areas where such resources occurred naturally. This, however, had to await, first, the accumulation of surpluses, human capital, and productive capacity accruing from the earlier stage of largely internal exchange and, second, the domestication of the donkey, which both enlarged the geographical horizon of southern elites and physically enabled them to engage in bulk export trade for the first time in their history.

We can only speculate about the historical consequences of these early patterns of trade, but I would suggest that their self-amplifying social ramifications would have created a situation in which the parallel development of multiple independent centers was a likely outcome. Such competition was surely a factor in the Sumerian Takeoff of the fourth millennium, and may well help explain why competing city-states continued to be the most characteristic political formation of alluvial Mesopotamia long after the end of the Uruk period. As already noted, political fragmentation and economic competition can promote accelerated social change. Patricia Crone (1989:161), for instance, argues that political fragmentation and inter-polity competition were crucial for what she perceives as the unique vitality of developmental rates in European polities of the late medieval and early modern eras as compared with those characteristic of other areas of the world at that time: "Far from being stultified by imperial government, Europe was to be propelled forward by constant competition between its component parts." Such may well have been also the case in ancient Mesopotamia.

In any event, in turning to trade both earlier and more intensively than elites in neighboring societies, elite individuals and institutions in Uruk Mesopotamia surely had no understanding of the long-term developmental consequences of the actions they were undertaking. Rather, trade simply became an efficient way to accomplish in the southern context what elites naturally want to do in all human societies, namely, sanction existing social inequalities, extend the amounts and varieties of commodities and labor at their disposal, and increase their political power at the expense of their rivals.

In this light, the Sumerian Takeoff became, in effect, an unanticipated consequence of long-term trade patterns that differentially favored the development of societies in the alluvial lowlands of Mesopotamia over polities in neighboring regions. This trade was inherently asymmetrical in its impact because, with some exceptions, it involved the import of raw or only partially modified resources from highland areas of the ancient Near East and the export of labor-added manufactured commodities from the southern Mesopotamian alluvium. At first, the trade was spurred by differences in productivity that favored the south and that were largely the result of geographical and environmental factors – what Cronon refers to as the “natural landscape.” Once a significant measure of exchange was in place, however, further conditions expanding and compounding the competitive advantage of Sumerian societies now arose mostly from the “created landscape” ensuing from the social ramifications of the trade. One such condition was provided by synergies derived from the greater density of population in rapidly urbanizing Uruk polities possessing ever-larger markets and pools of skilled and unskilled labor, usable, as needed, for commodity production, building or agricultural activities, as soldiers engaged in warfare against local rivals, or as colonists and emissaries sent to far away lands. In turn, synergies derived from
greater density and larger labor pools were compounded and expanded by the only sufficient conditions in the conjuncture: socially-created organizational efficiencies delivering ever increasing returns to scale from an ever more specialized labor force and allowing for exponentially more efficient and more accurate ways of conveying information across space and time. More than anything else, it is these social innovations that explain why complex regionally organized city-states emerged earlier in southern Iraq than elsewhere in the Near East, or the world.

BIBLIOGRAPHY

Adams, R. McC.

Adams, R. McC. and H.-J. Nissen

Algaze, G.
Amiet, P.


Bairoch, P.

Balinger, C.J.
2001 *City, Society, and State: The Role of Transport Costs in European State Development*. Master’s Thesis. Department of Political Science, University of North Carolina at Chapel Hill.

Benco, N.

Bourke, S.

Burghardt, A.F.

Carter, E.

Collins, P.

Cowgill, GL.
Crone, P.

Cronon, W.

Christaller, W.

Davis, S.J.M.

Delougaz, P. and H.J. Kantor.

Dercksen, J. G. ed.
1996 *The Old Assyrian Copper Trade in Anatolia*. Istanbul: Nederlands Historisch-Archaeologisch Instituut te Istanbul.

Emberling, G. and H. McDonald


Emberling, G.

Englund, R. K.


Erinç, S.

Farber, H.

Fernandez-Armesto, F.

Finkbeiner, U.

Finley, M.I.

Forest, J.-D.

Frangipane, M.

Fox, E.W.

Fujita, M., P. Krugman, and A.J. Venables

Fujita, M. and P. Krugman
Gibson, McG.  

Green, M.V.  

Goody, J.  

Hawley, A.  

Helms, M. W.  
1993 *Craft and the Kingly Ideal.* Austin: University of Texas Press.

Hicks, J.  

Hirth, K.  


Hole, F.  

Hudson, M.  

Jacobs, J.  

Jacobsen, T.  


Johnson, G.A.

Kolata A.

Kohl, P.

Kouchoukos, N.
1998 *Landscape and Social Change in Late Prehistoric Mesopotamia*. Ph.D. Dissertation, Department of Anthropology, Yale University.

Kouchoukos, N. and F. Hole

Krugman, P.

Lamberg-Karlovsky, C.C.

Larsen, M. T.


Leemans, W.F.


Lenski, G.


Levy, T.


Lösch, A.


Maekawa, K.


Malecki, E.J.,


Marcus, J.


Margueron, J. Cl.


Marshall, J. U.


McCormick, M.


McCorriston, J.


McNeill, W. H.


Mokyr, J.


Moo rey, P. R. S.


Mouck, T.


Myrdal, G.


Nissen, H.-J.


Nissen, H.-J., P. Damerow, and R. Englund

North, D.C.

North, D. C. and R. P. Thomas

O'Sullivan, A.

Oates, J.

Oates, J. and D.

Oppenheim, A. L.

Parrot, A.

Perrin de Brichambaut, G. and C. C. Wallén.

Petit-Maire, N, P. Sanlaville, and Z. W. Yan.

Pirenne, H.


Polanyi, K.


Pollock, S.


Postgate, J. N.


Potts, D. T.


Pournelle, J.


Powell, M.


Pred, A.


Ricardo, D.,


Renfrew, C. and J.F. Cherry, eds.


Rothman, M. ed.


Sahlins, M.


Sanlaville, P.


Schmandt-Besserat, D.


Schwartz, G.M.


Service, E. R.


Sherratt, A.

Silver, M.


Shennan, S.


Smith, A.


Smith, P.E.L. and T.C. Young, Jr.


Snooks, D.G.


Spufford, P.


Stein, G.


Steinkeller, P.


Stronach, D.

1994 Village to Metropolis: Nineveh and the Beginnings of Urbanism in Northern

Turkes, M.


Ur, J.


Vance, Jr., J.E.


Waetzoldt, H.


Warburton, D.


Weiss, H.


Wells, P. S.


Wheatley, P.


Wilkinson, T. J.


Wilkinson, T. J. and D. J. Tucker.


Wittfogel, K. A.


Wright, H.


Wright, H. T. and G. A. Johnson.


Wright, R.