THE SUMERIAN TAKEOFF*

GUILLERMO ALGAZE

ABSTRACT

The emergence of early cities in the alluvial lowlands of southern Mesopotamia during the fourth millennium BC must be understood in terms of both the unique ecological conditions affecting the area at the time, and the enduring geographical framework of the Mesopotamian lowlands which allowed for the efficient movement of commodities via water transport and facilitated interaction between diverse social units. These conditions promoted evolving long-term trade patterns that, inadvertently, differentially favored the development of polities in the southern Mesopotamian alluvium over contemporary societies in neighboring regions.

More specifically, by the final quarter of the fourth millennium, the social and economic multiplier effects of trade patterns that had been in place for centuries (if not millennia) had brought about substantial increases in population agglomeration throughout the southern alluvial lowlands. Concurrent with these increases, and partly as a result of them, important socioeconomic innovations started to appear in the increasingly urbanized polities of southern Mesopotamia. These were unachievable in

* Earlier drafts of this article were read and criticized by Robert McC. Adams (UCSD), Elise Auerbach (University of Chicago), Alex Joffe, David Jordan (UCSD), Philip Kohl (Wellesley College), Richard Meadows (Harvard), Joyce Marcus (University of Michigan), Joy McCorriston (Ohio State University), Joel Robbins (UCSD), Morris Silver (CCNY), Piotr Steinkeller (Harvard), and Donald Tuzin (UCSD). Each offered substantive criticism, important editorial comments, and crucial missing references. Their invaluable contributions are acknowledged with gratitude, even though, occasionally, I chose to disregard their advice. For this reason, remaining errors of omission and interpretation are entirely my own.

other areas of the ancient Near East, where urban grids of comparable scale and complexity did not exist at the time. Most salient among these innovations were (1) new forms of labor organization delivering economies of scale in the production of subsistence and industrial commodities to southern societies, and (2) the creation of new forms of record keeping in southern cities that were much more capable of conveying information across time and space than the simpler reckoning systems used by contemporary polities elsewhere.

More than any other factor, these socially created innovations help explain why complex, regionally organized city-states emerged earlier in southern Iraq than elsewhere in the Near East, or the world.

. . . without the wool for textiles to be traded for natural resources that were wholly lacking in the alluvium, it is difficult to believe that Mesopotamian civilization could have arisen as early and flourished as prodigiously as it did.

Adams 1981: 11)

INTRODUCTION

Robert McCormick Adams, as evidenced in the preceding quotation, always emphasized the critical role played by trade in the processes that gave rise to early Mesopotamian (Sumerian) civilization in the alluvial lowlands of the Tigris and Euphrates Rivers. In this article, I build on this theoretical foundation and try to elucidate, however roughly, the reasons behind the centrality of trade in the emergence of the world's earliest fully urban, state-level societies. My ideas owe much to a close reading of Adams's many seminal contributions on this and related subjects, as well to continued interaction with him over the years.

Economic geographers argue that the substantial variations in population concentration and economic activity present today are the inevitable result of cumulative processes whereby socially created technologies and institutions increase the natural advantages of specific sites or areas by delivering increasing returns to scale. This involves self-reinforcing processes of accumulation, exchange, population growth, agglomeration, and innovation 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).

These processes are not limited to the modern world, but instead have been operative at least from the time of the earliest cities and states. The emergence of early Sumerian civilization in the alluvial lowlands of southern Iraq during the Uruk period—roughly spanning the fourth millennium BC—is very much a case in point. This was characterized by the creation of a

thriving heartland of multiple competing but culturally unified city-states, a form of social organization that was hitherto unparalleled in the human career. Often singly and certainly in the aggregate, these emergent early Mesopotamian states greatly surpassed contemporary polities elsewhere in Southwest Asia in terms of their scale and their degree of internal differentiation, both social and economic (Adams 1981; Algaze 2001a; 2001b; Pollock 2001). A particularly noteworthy aspect of this process is that it followed centuries, if not millennia, in which the growth trajectory of polities in the alluvial lowlands of the Tigris-Euphrates system hardly differed from those elsewhere in the Near East (Oates 2001; Wilkinson 2001; H. T. Wright 2001a).

The crystallization of Sumerian civilization in the Uruk period thus represents a dramatic "takeoff"—a decisive shift in favor of southern Mesopotamia in the balance of early Near Eastern urbanization, sociopolitical complexity, and economic differentiation. 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 at this time that made it more probable that the takeoff would occur there rather than elsewhere? If the latter, what specific processes help account for the emergence of civilization in the south?

THE NEED FOR MODELS

Archaeological evidence for the emergence and growth of early cities in the southern Mesopotamian alluvium throughout the various phases of the Uruk period (ca. 3900–3200/3100 BC) is of varying reliability, resolution, and coverage[1]. The formative phases of the Early Uruk period (ca. 3900–3600 BC) are, for all practical purposes, known only through survey data (Nissen 1993). The Middle (ca. 3600–3400 BC) and Late Uruk (ca. 3400–3200/3100 BC) periods are somewhat better understood, since pertinent data are provided not only by settlement pattern surveys, but also by excavations at a few sites and some texts. However, existing exposures remain unrepresentative beyond the central administrative areas of major sites such as Warka, and available texts remain difficult to interpret (Englund 1998) and date only to the final stage of the Uruk period.

The scarcity of representative data bearing on the emergence and growth of early cities and states in fourth millennium southern Mesopotamia means that if we are to generate testable propositions about those processes, we need to use models illustrating how comparable phenomena unfolded elsewhere in

[1] For the chronology of the Uruk period, see now H. T. Wright and Rupley 2001.

the world. One obvious source is provided by the work of economic historians and geographers seeking to understand why, where, and how cities have emerged in modern times. These scholars take it as given that trade is a key factor in the evolution of social complexity and that cities serve as the most efficient way to manage regional and interregional exchange in situations marked by asymmetries in resource endowments, commodity production, and access to transportation across the landscape (for example, Hicks 1969; O'Sullivan 1996; Pirenne 1936).

In contrast, discussions in ancient Near Eastern studies about the why, where, and how big of early urbanism (when they occur at all) proceed as if Adam Smith's (1954 [1776]) work on the social ramifications of trade and David Ricardo's (1971 [1817]) analyses of the role of comparative advantage between regions as a spur to trade had never existed, and as if later refinements of their pioneering work by economic geographers and development economists (summarized in Krugman 1995) had never occurred. Particularly lacking in our field are discussions about how differences in access to transportation and the technologies of communication affect the process of urban growth. Worse still is the almost total absence in our discussions of other key causal agents highlighted in more recent economic analyses of urban dynamics. Without question, the most crucial of these is the role of urban institutions in creating organizational efficiencies that yield significant economies of scale in the production and distribution of commodities (North 1991).

The general lack of an economic focus in research on ancient Near Eastern urban origins is a direct consequence of the difficulty of characterizing, let alone quantifying, the data recovered from the earliest Mesopotamian cities. However, there are other complicating factors, both theoretical and methodological. On the theoretical front, many in our field assume, following the work of Karl Polanyi (1957a; 1957b) and Moses Finley (1985), that ancient socioeconomic phenomena were of a fundamentally different character from those in operation today, with the corollary that the forces underlying the emergence and growth of early cities in antiquity must also have been substantially different from those at work in more recent historical periods.

But is this really the case? I suspect that in many cases the answer is a clear no, and that this would be generally acknowledged if we only had the ability to precisely quantify and accurately characterize the types of economic activity in many urban societies of the premodern era. The historic periods of Mesopotamian civilization are a case in point. Polanyi's insistence, for example, that price-making markets subject to laws of supply and demand did not exist in early Mesopotamia 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 Isin-Larsa and Old Babylonian periods of the early second millennium BC (Farber 1978; Silver 1995). Equally relevant is recent work clarifying the entrepreneurial and profit-seeking nature of contemporary Mesopotamian trading enterprises. Cases in point include Old Assyrian trade with Anatolia (Adams 1974b; Dercksen, ed. 1996; Larsen 1976) and Old Babylonian trade in the Persian Gulf (Leemans 1960; A. L. Oppenheim 1960). Less well attested but possibly similar in structure were independent merchant colonies sent beyond Mesopotamia proper by individual Sumerian city-states in the second half of the third millennium (Foster 1993). The cumulative weight of this evidence makes it clear that wealth-maximizing behaviors by individuals, kin groups, and institutions played a role in spurring and maintaining Mesopotamian urbanism in the historical periods. Under these circumstances, can we truly discount the possibility that comparable motivations may also have played a role in the inception of Mesopotamian urbanism?

Another problem is methodological. Central Place theory, the tool most commonly used in approaching ancient Near Eastern urbanism, is inherently ill equipped to address questions of origin. In essence, the locational theories put forth by twentieth-century successors of von Thünen—most notably Walter Christaller (1966 [1933]) and August Lösch (1954 [1940])—seek to understand the forces that spread economic activity *away* from a center. This involves a 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. Central Place models help us understand how hierarchies of function are maintained across a landscape. However, because the models simply assume the a priori existence of a central urban market, they tell us nothing about why population and economic activity become concentrated *in the first place*, as noted by John Marshall (1989) and Paul Krugman (1995).

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. From a Ricardian perspective, cities invariably represent nodes in wider transportation networks and are the most efficient way to mediate sustained trade between places with varying degrees of comparative advantage in the production of both necessary and desirable resources. Such advantage is created by differences in productivity between polities, caused by the naturally uneven distribution of resources in different areas, differences in access to transportation, and social factors such

as organizational and technological efficiencies delivering increasing returns to scale.

Since cities form as a response to regional imbalances in comparative advantage, they should preferentially emerge either at natural passage points between regions involved in exchange or at the end points of natural transportation routes between such regions (Burghardt 1971; Hirth 1978). They should also form at critical nodes along such routes, such as intersections or transshipment points where goods need to moved between routes served by different types of transport—for example, from water to land transport (Bairoch 1988; Burghardt 1979). Further, because of the multiplier effects of trade on social evolution (below), cities should concentrate in areas possessing the greatest positive productivity differential along a transportation route which in turn ensures 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 by the iconoclastic urban expert Jane Jacobs (1969; 2000). 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 first created as a result of positive feedback created 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 possible the acquisition of more and different imports, 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 (that is, human capital), and this, in turn, creates the potential for further economic diversification by adding new types of work and new ways of working. 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 for understanding how cities grow is that of Circular and Cumulative Causation. This is an idea first articulated in the 1950s by the economist Gunnar Myrdal (1957) and later elaborated, expanded, and formalized by Allen Pred (1966). At its simplest, it 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 that are amenable to economies of scale, with, for instance, mass production of local manufactures. This substitution will further expand urban employment, drawing in workers from the countryside and other regions, and in so doing will also increase the size of the local market and the range of skills possessed by the population. As this process unfolds, a number of multiplier effects come into play, resulting in increases in productive capacity. These include the creation of new industries that provide production inputs to the initial industry ("backward linkages") or of activities that either add further value to semi-finished goods or provide services connected to, and made necessary by, the production of such goods ("forward linkages") (Pred 1966: 25–26).

In the Mesopotamian case, one of the most important of these forward linkages must have been the expansion of the managerial classes required to organize the larger number of workers, store and distribute their enhanced production, and keep accurate records of these various activities. This expansion is likely to have been quite substantial, as may be inferred from recent studies of the relationship between changes in population density and the extent of bureaucratic superstructures in modern urban societies. These studies show that expanding population density consistently leads to a disproportionately large growth in the size of the communicative components of managerial institutions in a city. Why this is the case is explained by John Kasarda (1974), a sociologist, who notes that in human societies, as in biological organisms, increasing size exacerbates particular systemic problems and normally results in disproportionate growth in sectors serving to solve these problems. The most critical problem faced by large-scale social systems, according to Kasarda, is articulating communication among its parts. For this reason, as they grow increasingly large and diverse, complex societies must divert an ever larger proportion of their human resources to collecting, processing, and transmitting information.

In due course, these interrelated multipliers combine to create the enlarged population and market size necessary to induce more import substitution. In turn, as the process is repeated at an ever increasing scale, a 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 increase 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 CITIES: 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 to mind. First, as they have in modern times, trade and changes in commodity production and labor organization are likely to have played fundamental roles in change. 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 gained an initial advantage, those processes would have concentrated new growth and its multiplier effects in the already expanding region rather than elsewhere (Malecki 1997: 49–50). Under these circumstances, what is left to be elucidated are the forces that set trade (and its multiplying ramifications) into motion in the first place.

In the case of early Sumerian civilization, the trigger was provided by the combination of the enduring geographical framework of the Mesopotamian alluvium and the unique environmental conditions that prevailed in the area during the late fifth and throughout the fourth millennium BC. Geography was important because it provided southern Mesopotamian societies with an enduring and irrevocable advantage over their neighbors in the form of the low shipment costs made possible by water transport. The cities of the alluvium were, in effect, at the head of an enormous dendritic transportation system created by the north–south flowing 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 the southern cities lay in their ability to import needed commodities *in bulk* from resource areas in the surrounding highlands by means of low-cost river transport by raft or boat. 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.

Environment, in turn, also gave southern societies important advantages in the productivity, variety, and resilience of their subsistence and exportable resources as compared with those available to surrounding polities. These advantages would have been particularly pronounced at the time of early urban emergence in Mesopotamia. As I have recently discussed these data elsewhere in some detail (Algaze 2001a), only a brief summary of the pertinent information is offered here.

A variety of paleoenvironmental data suggests that the climate was wetter for much of the fourth millennium BC. This means that marginal areas of the Mesopotamian alluvium that are today unproductive because of insufficient water or lack of drainage would likely have been integrated into fluvial networks draining into the sea. In addition, at least for the Late ʿUbaid period and, possibly, into the Early Uruk period, parts of the alluvium which today receive no summer precipitation whatsoever would have been affected by summer monsoonal rains. These rains would have greatly expanded the availability of animal forage at precisely the time of greatest need, and would have enhanced the productivity of both summer crops and date palm groves.

Also more propitious for human settlement were the regimes of the Tigris and Euphrates Rivers through the fourth millennium. New remote sensing work at UCSD by Robert McC. Adams and Jennifer Pournelle (this volume) shows that the rivers formed a single dynamic and complexly intertwined network of anastomosing channels throughout the fourth millennium, creating much vaster areas than was the case later on when various types of high-value crops and grain could only have been produced by means of simple flood-recession irrigation.

Last but not least was the enhanced availability in the fourth millennium of resources from bio-mass-rich marshes, lagoons, and estuaries brought into close proximity to Uruk population centers by temporary mid-Holocene sea level rises. These resources included dried, salted, and smoked fish, various types of fowl, numerous reed products, and dairy products from herds of cattle kept at the margins of greatly enlarged marshes and lagoons (Pournelle, this volume).

The consequences of the geographical and environmental advantages just noted are clear. On the one hand, advantages in the 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 would have allowed Uruk elites to mobilize surpluses from their immediately dependent hinterlands at lower cost than their competitors. This also meant that the extent of those dependent hinterlands would naturally be larger than those of their landlocked competitors, and that southern elites and institutions could procure resources and information from a much vaster area at much lower cost than their contemporaries. Taken together, these advantages gave emerging Uruk polities substantial comparative advantages 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 Ricardo reminds us, trade is the logical outcome.

EARLY MESOPOTAMIAN CITIES: WHERE?

It should be clear from the foregoing discussion that many of the key underlying assumptions of Central Place theory are simply unsuitable to the analysis of the conditions prevalent in the alluvial lowlands of the Tigris-Euphrates fluvial system when the early cities first emerged. First, the model assumes that access to agricultural land is the most important economic variable affecting the location of early cities and, therefore, that the movement of agricultural products is the main factor structuring the spatial relationship between settlements in any given region. Second, the model assumes that transportation costs increase steadily with distance. These assumptions are perfectly compatible with conditions prevalent across much of the Mesopotamian periphery, but they fail to account for the complexity of the situation in southern Mesopotamia. The first assumption misses the mark in that it underestimates both the degree of ecological variability of the southern alluvial plain in the fourth millennium BC and the importance of marshes, lagoons, and estuaries at the time. Further, it discounts the probability that resilience strategies based on access to multiple resources (Adams 1978) within the alluvium or the desire to maximize access to trade and communication routes in and out of the alluvium (Algaze 1993a) were factors of equal or even greater locational importance for early cities in the Mesopotamian alluvium than access to agricultural resources. The second assumption has always been irrelevant to conditions in southern Mesopotamia because transportation costs for bulky, locally produced agricultural commodities would not have increased exponentially with distance within the alluvial delta of the Tigris-Euphrates system as presumed by Central Place models. Rather, as explained earlier, such costs were kept in check by networks of natural and artificial canals and marshes surrounding early settlements.

Available surveys of the Mesopotamian alluvium show that multiple urban centers emerged in the area during the fourth millennium (Adams 1966; 1981; Adams and Nissen 1972; Gibson 1972; H. T. Wright 1981b). Each of the sites in question is located along one of the main courses of the Tigris and Euphrates Rivers as they existed at the time. Access to (irrigable) agricultural land was no doubt a factor here, but, above all, the large Uruk sites (Nippur, Uqair, Kish (?), Abu Salabikh, Site 1306, and Site 1172) of the northern part of the Mesopotamian alluvium were also located along the natural transportation routes in and out of the alluvium, which follow the rivers. The southernmost Uruk-period centers (Uruk, Umma, Eridu, Ur, Girsu [Tello], Site 125), in turn, were all situated at logistical transshipment points between different types of routes. They were (1) at the end of the vast north–south transport route created by the rivers, (2) within reach of the lateral routes created by extensive marshes and lagoons at the encroaching head of the Persian Gulf in

the fourth millennium (Pournelle, this volume), and (3) at the head of overland routes into the Arabian Peninsula and maritime routes into the Persian Gulf and Indian Ocean. These uniquely privileged positions naturally lessened the transportation costs of these sites and maximized their access to information and products from various far-flung areas and catchments.

EARLY MESOPOTAMIAN CITIES: HOW?

The concept of Circular and Cumulative Causation, discussed earlier, allows us to visualize a still speculative though ultimately testable scenario to account for the precocious urban takeoff of southern Mesopotamian societies in the fourth millennium. For heuristic purposes, this evolving process is divided here into a number of discrete stages, although substantial overlaps clearly existed between them.

The initial stage in the growth of southern economies would have taken place during the late fifth and early fourth millennia—a time when the Mesopotamian alluvium was a mosaic of very different but easily exploited resource areas. 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 communities exploiting these varied economic resources. Each of these centers would have specialized in the production of those few crops or commodities best suited 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 communities situated at the margins of the better-watered parts of the alluvium where they would have enjoyed preferential access to pastoral and nomadic groups; (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 centers immediately by 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 already in place. In this stage, technologies and practices initially developed by individual centers exploiting their own specialized niches would have come to be perceived as highly advantageous by many of the competing centers. This naturally would lead to a decrease in regional specialization within the alluvium, as each competing

polity used the material surpluses and human skills acquired during the initial 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 import-substitution (see above).

The third stage of the process, datable to the Middle and Late Uruk periods (ca. 3600–3200/3100 BC) 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 in external trade between individual cities and neighboring areas. Here, the ongoing import substitution processes in the south would have begun to focus more on the replacement of foreign commodities. Increased foreign trade was also enhanced at this point by the domestication of donkeys, which made overland communication across the ancient Near East possible (H. T. Wright 2001a: 127), and likewise, for the first time, provided the physical capacity to export alluvial goods in bulk.

As external trade became increasingly important in the Middle and Late Uruk periods, various types of southern outposts were established at locations strategic for transport across the Mesopotamian periphery (Algaze 1993a; 2001b). These outposts were principally, but not solely, situated at the intersections of the north–south flowing rivers and the principal east–west overland routes across the high plains of northern Mesopotamia (Algaze 1993a; 2001b; but see Rothman, ed. 2001). While these outposts may have served in part as outlets for displaced populations from the south (Johnson 1988–89; Pollock 1999; Schwartz 2001; H. T. Wright 2001a), their carefully selected locations suggest that they also served as collection and transshipment points for the increasing amounts of commodities imported into the alluvium in the later part of the Uruk period, and as distribution points for alluvial exports.

The role played by these trade patterns in the emergence of Sumerian civilization is made clear if we focus our attention on the long-term multiplier effects of import substitution. These processes can be easily documented in the archaeological record of ʿUbaid- and Uruk-period Mesopotamian societies. Perhaps one of the earliest examples in the record is the partial replacement of imported flint by locally manufactured clay sickles, a process that began in the Late ʿUbaid period and continued through the various phases of the Uruk (Benco 1992). Another example is provided by metals, first attested in the south by the end of the ʿUbaid (Moorey 1994: 221, 255–258). Initially, metal goods must have been brought into Southern Mesopotamia as fully finished products imported from the metal-producing highland regions of Iran and Anatolia where metallurgical technologies were first developed (Kohl

1987: 16; G. 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 this 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). In addition, by the final phase of the Uruk period, we also get textual corroboration for the shift in the form of the pictogram used in the earliest Archaic Texts to denote a smith, which shows a smelting furnace with attached blowpipes (Moorey 1994: 243).

By far the most important case of import substitution in the south is provided by the adoption, sometime by the Late ʿUbaid, of wool-bearing breeds of sheep that had been 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 sites show that by the second half of the fourth millennium, these once 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 which, 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 that time.

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 may 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 into their new man-made habitat. The second is that the south had the comparative advantage of easy access to many natural dyes. This is another point recently raised by McCorriston (1999; 2001: 222, and personal communication 2001), who notes that many of the dyes used to color Near Eastern wool in antiquity could be derived from desert or garden plants available in or around southern Mesopotamia or from products only obtainable via the Persian Gulf, such as various types of marine gastropods and indigo.

The third factor is that southern Mesopotamian societies possessed larger pools of labor available for textile work. From the beginning, these workers appear to have been organized in ways that allowed for greater efficiency and superior craftsmanship in the production of textiles. Contemporary depictions of female workers (pig-tailed figures) attending horizontal looms (Amiet 1972: nos. 673–674; 1980: nos. 319–320) suggest that already by the Late Uruk period the woolen textile industry of the southern lowlands was based on centrally administered weaving establishments exploiting the labor of various categories of dependent women in both cities and the countryside, such as we know existed in most Sumerian and Babylonian polities during the third and early second millennia BC (Jacobsen 1970 [1953]; Maekawa 1980; Waetzoldt 1972).

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. Particularly noteworthy are 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 written record of the later third millennium, and both require a substantial input of value-adding labor and new resources (McCorriston 2001: 222; Potts 1997: 95). Under Mesopotamian conditions, an equally important forward linkage is the shift to industrial-scale textile production which would have required scores of bureaucrats to record, store, and redistribute the output, and also to supervise the distribution of subsistence rations to workers. 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 include pasturing the sheep, washing, plucking and/or shearing, combing, and spinning the wool, separating the wool by quality, and delivering it to the various urban and rural locations where state-organized weavers labored.

THE EVIDENCE FOR TRADE

The foregoing discussions presume that trade, both internal and external, was the engine of early Mesopotamian urban growth. Nevertheless, substantial disagreement remains on the importance of long-distance trade to the processes of urban and state formation in southern Mesopotamia.

Many scholars reviewing the data for southern Mesopotamian economies of the fourth millennium properly highlight the importance of local tribute extraction and intraregional distribution of resources as key elements in that economy, but either minimize the overall importance of long-distance

trade to the socioeconomic processes at work at the time (for example, Frangipane 2001a; Pollock 1999; Schwartz 2001; Weiss 1989) or presume that the increase in long-distance exchange was a consequence of urbanism rather than a cause (H. T. Wright 1981a; 2001a). 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 to metals which must have been imported into the south either as finished products or, most likely at this point, manufactured in place from imported ores and ingots (above). Similarly, some of the Middle and Late Uruk–period southern colonial sites on the Euphrates yielded evidence for the import and in situ processing of various types of exotic stones, including lapis (for specific references, see Algaze 2001a: 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. 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 a full inventory, see now Forest 1999: 67–73).

The second problem is that those who minimize the early importance of trade consistently overestimate how representative the archaeological record really is. Must we believe that the Riemchengebäude was an exceptional find that bears no relationship to elite activities at the site? On the contrary, it is more parsimonious to think that this is not a case of unusual wealth but rather one of unique preservation, since the building was buried after being consumed by fire. Most likely, one reason why wealth on the scale found at the Riemchengebäude is not well attested in other excavated areas of Warka, save for scattered buried hoards such as the Sammelfund (Heinrich 1936), is because so many of the Uruk-period structures cleared in the Eanna Precinct represent no more than the foundations of buildings that were carefully and purposefully cleaned and emptied in antiquity (Eichmann 1989).

An equally important reason for the paucity of evidence for Uruk-period exchange is that so many of the articles traded at the time would have left few traces in the archaeological record. This includes, of course, both the principal exports (finished textiles) from the alluvial lowlands at this time and the principal imports, either because they do not preserve (timber), or because, except in destruction levels, they are commonly recycled (metals, precious stones, and the like).

Andrew Sherratt (2004) also notes that in truly complex societies, sumptuary goods 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 (for example, metals by melting), and that they will be passed on across generations more consistently. As commodities circulate more broadly across wider social networks, excavations at single central sites, or, worse still, at the core of such sites, would be unlikely to produce a representative sample of the scale and type of sumptuary commodities in circulation at any one time. Regretfully, the kind of broad sampling strategy needed to document the spread of sumptuary commodities in regionally integrated societies has not been part of the excavated record for the Uruk period in the alluvium.

A further problem is that in complex societies, a high proportion of exotics would also naturally get withdrawn from circulation for use as burial gifts. One suspects that the continuing and still puzzling dearth of mortuary evidence from southern Mesopotamia for the Uruk period may be responsible to a greater degree than is generally acknowledged for the view that imports of sumptuary commodities into fourth-millennium Mesopotamia were relatively rare.

ISSUES OF SCALE: SOUTHERN VERSUS NORTHERN MESOPOTAMIA

Where trade flows, its ramifications soon follow in the form of increasing social complexity and urbanism. 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 patterns of settlement in the alluvium against comparable data from neighboring regions, particularly from upper Mesopotamia.

For the south, available survey data (Adams 1981; Adams and Nissen 1972; H. T. 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 they were in any one coherent area of the Mesopotamian periphery (Kouchoukos 1998: tables 5.4–5.6, fig. 5.9).[2] In

[2] 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 largely 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 their northern counterparts.

fact, surveys document multiple interacting urban sites (40+ hectares in extent) 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. Development in the area reaches its peak by the final phase of the Uruk period, when the site of Warka grew to the extraordinary size of 250 hectares (Finkbeiner 1991) and was surrounded by numerous dependent villages and towns, totaling a minimum of 280 hectares of further occupation (Adams 1981; Adams and Nissen 1972).

The sites just discussed are likely to be only the tip, so to say, of the Uruk-period settlement iceberg in southern Mesopotamia because a number of sites exist outside of the thus far surveyed areas 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 (for instance, 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.[3] Numerous Archaic tablets recently plundered from either (or both) of those sites appear immediately comparable with the earliest examples from Warka (Robert Englund, personal communication 2001). At a minimum, these tablets attest to the economic importance of the Umma area in the Late Uruk period, but since at Warka 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 may have been second only to Warka itself in terms of urban and social development in the Late Uruk period.[4]

The long sequence of growth in the south throughout the Uruk period contrasts starkly with the overall developmental trajectory of contemporary

---

[3] Both Umma (WS 197) and Aqarib (WS 198) were at the edge of Adams's 1968 survey area but could not be properly surveyed at that time because of extensive sand dunes covering both sites (Adams and Nissen 1972: 227–228). The dunes have since cleared the area.

[4] At present, the settlement data for Late Uruk southern Mesopotamia is anomalous in that it shows a central site (Uruk) that is four times as large (that is, populous) as second-tier settlements (such as Site 1306). Analyses of modern urban systems (Krugman 1996) leave little doubt that urban populations arrange themselves in rank order by size in ways that are both patterned and predictable ("Zipf's Law"). The behavior of such modern systems suggests that a tier of settlement may actually be missing from our southern Mesopotamian data. If urban rank-size rules apply to the Mesopotamian case, as I would expect, the missing settlement(s) should be roughly about half the extent (population) of Warka. I expect that Umma represents the missing tier and that further work at that site will eventually show it to have been somewhere in the range of 120 hectares in the Late Uruk period.

northern Mesopotamian societies. To be sure, as Henry Wright (2001a: 145) has noted, both have an initial burst of settlement growth and expansion of social complexity in the earlier part of the fourth millennium, but in the north this lasts only until about 3500/3400 BC. This has become evident only recently in upper Mesopotamia as a result of new excavations at Tell Brak (J. Oates and Oates 1997; Emberling et al. 1999) 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 both Tell el-Hawa (Wilkinson and Tucker 1995) and Tell Hamoukar (Gibson 2000; Gibson et al. 2002; Ur 2002a; 2002b; Reichel 2002) and their environs in the Jebel Sinjar plains of northern Iraq and Syria, and older surveys of Samsat and its environs (summarized in Algaze 1999) in the upper Euphrates area of southeastern Turkey. This new 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 65 hectares (Emberling et al. 1999) in the so-called northern Middle Uruk period and may have been even larger at the time, depending on whether or not the intervening area between the main site and nearby contemporary suburbs was occupied (H. Wright, personal communication 2004; J. Ur, personal communication, 2003). Brak was thus broadly similar in extent, if not slightly 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-hectare range. Hawa is reported to have been in the 30+-hectare range at this time, and Samsat and Hamoukar were about half that size.

Similarities in the scale of individual sites in northern and southern Mesopotamia, however, mask important differences in the complexity of the overall settlement systems of both areas. Even at their peak in the "northern Middle Uruk" period, these Late Chalcolithic societies hardly equaled their southern counterparts. This is reflected in both the density and the hierarchy of settlement grids surrounding large settlements in the two areas. Pending the publication of recent surveys around Brak conducted by H. Wright and his colleagues (Lawler 2006), the best data we have for the north is 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 1990a; 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/Mid-

dle 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).

A further noteworthy difference is a consequence of the geographical constraints affecting human settlement in northern and southern Mesopotamia. Large Late Chalcolithic settlements in the northern plains such as Nineveh, Hawa, Brak, and Samsat were situated in different drainage basins and were separated from each other by hundreds of kilometers. They were thus largely isolated one from the other 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 one another.

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 and Hamoukar show that those settlements contracted in the second half of the fourth millennium (Emberling 2002; Emberling et al. 1999: 25–26; Gibson 2000; Gibson et al. 2002; Ur 2002a; 2002b), just as the expansion of southern sites such as Warka reached their Late Uruk peak. Brak was the largest and most impressive of the Late Chalcolithic centers in northern Mesopotamia identified thus far, and its contraction in the final quarter of the fourth millennium meant that Late Uruk urban centers in the alluvium were significantly more developed than contemporary Late Chalcolithic polities in the Mesopotamian periphery. In fact, at 250 hectares, Late Uruk Warka is likely to have been exponentially 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.

Survey evidence from various areas across northern Mesopotamia (summarized in Algaze 1999: table 3; Wilkinson 2000) shows that the area was effectively ruralized by the end of the fourth and the transition to the third millennium. The indigenous centers that existed in the region in the preceding period had vanished, and comparable indigenous centers did not reappear in the area until the final phases of the Ninevite V period, sometime in the second quarter of the third millennium (Schwartz 1994a; Weiss 1990; Wilkinson 1994). In contrast, urbanism in the southern alluvial plains continued to flourish and expand not only though the Late Uruk period, but

throughout the fourth–third millennium transition (Jemdet Nasr) as well (Adams 1981; Postgate 1986). The urban spiral of the south continued unabated well into the first quarter of the third millennium (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; Gibson 1972; H. T. Wright 1981b). Warka reached 600 hectares 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 (E. Carter 1985).

THE SYNERGIES OF CIVILIZATION

Multiple socio-evolutionary synergies would have arisen from the differences in rates of population agglomeration and the distance between polities typical of southern Mesopotamia and areas on its periphery throughout the second half of the fourth millennium, outlined above.

The first synergy arises from the greater concentration of polities that existed in the Mesopotamian alluvium throughout the roughly 700-year duration of the Uruk period, as compared with 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. The impact of these mutually reinforcing processes has been explained by Robert Wright (2000: 165–168), 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, 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 hinterlands through the Uruk period as compared with the more dispersed settlement typical for neighboring areas at the time. This had many consequences. First, as Adam Smith (1954 [1776]: book 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). In addition, 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 or need.

Increasing population density in towns and cities would also have compounded the natural transportation advantages of the alluvial environment by the development of further efficiencies in shipping and communication arising from the increasingly compact arrangement of the inhabitants of the area throughout the fourth millennium. One compounding efficiency was provided when artificial canals began to be built across portions of the southern alluvium in the Uruk period (Adams 1981; Pournelle, this volume), expanding the communication network to areas beyond the natural flow of the rivers. In so doing, the new canals also served to reinforce ongoing urbanization processes in the alluvium. This effect, no doubt inadvertent, is suggested by studies that clearly link reductions in transport costs of agricultural commodities in traditional societies to the movement of population into cities (Fujita and Krugman 1995: 520).

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 cities. As interactions multiply, information flow is enhanced. This has two crucial interrelated effects. The first is that it exponentially increases the possibility that technological improvements and inventions will take place in urban settings. 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 rises as the store of information increases. This increase will always be logarithmic, since the possibility of new combinations increases many times faster than the basic elements of information. The second effect of enhanced information flow is one noted earlier: innovations, once created, would naturally diffuse faster.

As the web of interpersonal communications became increasingly dense in the southern cities that were growing many times larger than neighboring population centers by the second half of the fourth millennium, 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 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 technologies of the intellect was the *systematic* use of various types of dependent laborers receiving rations for the production of subsistence and sumptuary commodities and for building activities. Evidence for this is provided by millions of ration containers themselves found at a variety of Uruk sites (Nissen 1988), by hundreds of attestations of signs for various categories of dependent laborers in the known corpus of Archaic Texts

from Warka, and by at least one tablet that summarizes food rations given to different groups of male and female captives (Englund 1998: 70, 178–179, fig. 66). This means that elites in fast-growing Uruk cities 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 also 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 in how labor was organized is provided by the well-documented shift to standardized, mass-produced ceramics throughout the Uruk (Nissen 1974), but comparable changes can be seen in the way other commodities, such as wool (Green 1980; Nissen 1986b) and metals (Nissen 1988), were produced or procured at that time. 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 technology of the intellect 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. This process started in earnest in the Middle Uruk period with the introduction of devices such as impressed hollow balls filled with tokens capable of conveying information by combining numbers and images, and culminated in the development of pictographic writing by the final phase of the Uruk period. Writing was arguably the most important consequence of the disproportionate expansion of the communicative sectors in growing Mesopotamian cities which resulted from their need to articulate their ever more diverse and ever larger components. What writing gave to early Mesopotamian decision makers (and the urban institutions they worked for) was a flow of varied and reliable economic information that allowed them to deploy available labor and resources so as to maximize their revenues and extend their power. In so doing, writing provided southern Mesopotamian societies of the Late Uruk period with a further substantial competitive advantage over contemporary polities elsewhere in which similar breakthroughs in accounting, accountability, classification, and access to information were absent (Algaze 2001a).

These various synergies represent multiple facets of a single phenomenon: advances in the efficiency and intensity of social interactions possible within and between southern 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, human settlements have historically 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 has been understood by economists since the time of Adam Smith (1954 [1776]: book I, chaps. 1–3), who observed that gains in the efficiency of transportation and communication always act as a spur to economic 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 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 all of this relates to the Sumerian case. The Uruk-period takeoff correlates both with enhanced communication efficiency in the form of new reckoning and writing systems and with reductions in mobility costs as population became concentrated, production facilities consolidated, and production standardized.

CONCLUSIONS: THE MESOPOTAMIAN CONJUNCTURE

Why did the balance of urbanization and social complexity in the ancient Near East shift decisively to the southern alluvial lowlands of Mesopotamia in the second half of the fourth millennium BC? Early on, the stage was set by advantages in productivity, reliability, and ease of transportation inherent to the "natural landscape" of southern Mesopotamia. Absent in neighboring regions, these advantages can be considered as the necessary conditions in the conjuncture that resulted in urban development. No doubt, the most important of these advantages was ease of transport. Whereas geography in the south both permitted and encouraged linearly arranged agglomerations based on intensive agriculture and on boat and raft transport, the geography of areas outside the Tigris-Euphrates alluvial lowlands instead encouraged population dispersal so as to maximize the amount of territory under (dry farmed) cultivation. 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 in neighboring areas in the fourth millennium. Accordingly, polities in those areas were not likely to significantly enhance

their productivity in and of themselves, 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.

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 these upper Mesopotamian polities were able to marshal the organizational efficiencies needed to overcome the difficulties of overland transportation that had prevented their Late Chalcolithic predecessors from forming enduring complex societies. Bluntly put, urbanism in the Mesopotamian periphery was only possible as an engineered landscape; it only became viable as a result of innovations in communication and labor control created first in the Sumerian heartland.

In the end, it turns out that Karl 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 as in 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, 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" (Fernández-Armesto 2001: 182).

Nevertheless, natural advantages derived from geography and environment do not explain in and of themselves the crystallization of early Mesopotamian civilization—or, for that matter, that of any other pristine civilization. In the final analysis, environmental and geographical factors are only permissive, not prescriptive. Whether individuals and groups react to environmental conditions and take advantage of geographical possibilities, and how they do so, are always constrained by culturally determined percep-

tions of opportunities and threats. Moreover, the present is also shaped by the past through inherently unpredictable accidents and innovations that add an element of indeterminacy to any attempt at historical prognostication or explanation.

Seen in this light, the natural advantages of the southern Mesopotamian landscape merely provided a backdrop wherein some social responses became more likely than others. Given the diversified but dispersed resources prevalent in southern Mesopotamia throughout the late fifth and fourth millennia BC, and the naturally low mobility costs of the area, one of the most probable of these responses was for pre- and proto-historic elites to specialize in the production of a limited number of commodities best suited to their location within the alluvial environmental mosaic and to engage in trade with differently specialized local rivals. By the same token, the absence of important resources from the Mesopotamian environment, most notably roofing-grade timber, stone, and metals, also made it likely that early southern elites would seek to engage in trade with their foreign counterparts in areas where such resources occurred naturally.

In turning to trade earlier and more intensively than those in neighboring societies, elites in alluvial 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 what elites 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. In this light, the Sumerian takeoff was, 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.

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. Once a significant measure of exchange was in place, however, conditions that further expanded the comparative advantage of Sumerian societies arose primarily as a result of conditions created in large part from the social ramifications of the trade itself. Rapidly urbanizing Uruk polities possessed 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 faraway lands. Synergies derived from the greater population density and larger labor pools were compounded by socially created organizational efficiencies delivering ever increasing returns to scale. These came principally in the form of new ways of organizing labor and in more efficient and more accurate ways of conveying information across space and time that southern

societies developed through the Middle and Late Uruk periods, culminating in a formal writing system. Social innovations such as these ultimately explain why complex, regionally organized city-states emerged earlier in southern Iraq than elsewhere in the Near East, or the world.