Is Urbanization in the Global South Fundamentally Different?
Comparative Global Urban Analysis for the 21st Century

Gregory Randolph, University of Southern California (USC)¹
Michael Storper, London School of Economics (LSE) and University of California, Los Angeles (UCLA)

This version: November 2021
Accepted and forthcoming in Urban Studies (2022)

Abstract

A vigorous debate has emerged in recent years over how to understand cities of the Global South. A pivotal issue in this debate is whether urbanization processes in the South are so fundamentally different from historical and current urbanization in the Global North that many of the theories developed from studying the latter have limited utility in application to the former. In this paper, we review evidence from a range of disciplines on recent and ongoing urban transitions and the dynamics of urbanization in the Global South, attending to features that distinguish the urban South from the urban North. Our reading of the evidence indicates that parts of the Global South may be urbanizing along historically and geographically specific trajectories; however, we argue that these differences are best understood through a unified set of global urban theories. Rather than flattening or silencing difference, theories that seek generalization across time and space sharpen the identification

¹ Corresponding author: grandolp@usc.edu
and appreciation of key differences in urbanization processes. We can generate context-sensitive explanations of urbanization and context-appropriate policy advice by analyzing how the fundamental dynamics of urbanization recombine and interact with one another as they are interwoven with a host of local specificities. But likewise, we can also generate a sense of policy challenges that cut across many cities, both within and between the Global North and South, by using the tools of social science to do so.

1. The challenge of understanding and explaining Global South urbanization

Are urbanization processes in the contemporary Global South different in their nature and genesis from their counterparts in the Global North, either past or present? How should theories developed from studying urbanization in the North be reevaluated in light of urban transitions in the South, or theories developed from the perspective of the South be re-evaluated in light of comparisons to historical processes in the North? A spirited debate around these questions has emerged in recent years. Some scholars argue that canonical urban theories are limited by their geographical origin in Northern cities and the limited experience of Northern scholars (Robinson, 2002; Roy, 2009), circumscribing their utility in application to the South, especially given empirical differences in the urban South such as the pace of urbanization and the legacies of colonialism (Sheppard, 2014). For many of these scholars, examining the urban South with a fresh lens is not about constructing a better, unified urban theory; it is about resisting attempts to universalize (Roy, 2016) by “provincializing” older urban theories (Sheppard et al., 2013) and pursuing “mid-level” theory located somewhere between particularism and universalism (Ghertner, 2015; Roy, 2005).
While these perspectives do not constitute a monolithic or unified critique (Lawhon and Truelove, 2020), they contrast with a long history of treating urbanization as a global historical process about which social scientists make theoretical claims across space and time (Bairoch, 1988; Braudel, 1992; Childe, 1950; Davis, 1965; Dyson, 2011; Fox, 2012). Scott and Storper (2015) maintain that, despite empirical differences between cities of the South and North, core concepts from established urban theory remain relevant in understanding the former, and that the basic tenets at the core of urban and spatial economics are useful in understanding the South. Mabin (2014) argues that simply including more geographies in the frame of inquiry does not, on its own, produce new theories.

The field of urban studies has always struggled to account for the complexities of urbanization. The current debate updates this long-standing challenge in three ways. First, the Global South is where most of the world’s urban growth has been taking place over the past 70 years; by turning its attention to the South, the field of urban studies is expanding the evidentiary base upon which our knowledge of urbanization is built. Second, this wider evidentiary base allows us to ask whether inherited theories of urbanization pose the right questions in the first place or whether established research agendas are provincial. Third, understanding the nature of similarity and difference among cities across space and time is critical to considering policies appropriate to the cities of today’s Global South. These motivations for studying the urban South are found in different disciplines. In urban economics, for instance, Chauvin et al. (2017) are concerned with “a knowledge mismatch, for urban economists have predominantly focused on the cities of the wealthy West” (17). In urban planning, Watson (2009) argues that the “techno-managerial” (2259) approach to planning in the North is not appropriate in settings where most people live in conditions of informality.
We deliberately use a specific definition of the term *theory* to mean a proposition describing a structural relationship between at least two variables. Most theories expect significant empirical variation in this relationship, expressed as a distribution, probability, or elasticity, all of which can be the result of local factors that differ from place to place. Social science theories almost never approach the status of *laws*, in which the relationship between two variables is so strong that extremely little empirical variation is observed. Theoretical determinants are therefore not deterministic. It follows that no single urban theory—regardless of where it is deployed—is expected to provide a complete *explanation* for any real-world urban phenomenon. Rather, it supplies necessary analytical scaffolding that can be used, alongside other types of information, including additional theoretical relationships and geographical and historical specificities that interact with them, to construct an explanatory narrative. To use an analogy, a theory of climate does not on its own explain the weather, but nor can weather be explained without a theory of climate.

In this paper, we examine evidence on Global South urbanization through a South-North comparative lens. We apply theories of general relationships between urbanization and four variables—demography, economic development, migration, and the production of the built environment. By urbanization, we mean a process by which societies increasingly organize themselves into physically and demographically dense settlement patterns that function through a complex division of labor. The four variables whose relationship to urbanization we analyze are not the only dimensions on which the urban South and North could be compared. However, we use them to illustrate how theories of urbanization, by attempting to generalize across time and space, can effectively illuminate key *differences* between South and North.
Before proceeding, what is the Global South? Scholars offer three different kinds of definitions: a world-regional classification, with the South constituted by Africa, Asia and Latin America (Mitlin and Satterthwaite, 2013); a “concept-metaphor” definition, where the South represents places of marginality and dispossession anywhere in the world (Sparke, 2007); and a definition in which the South denotes countries or regions where the majority of people face poverty and vulnerability (Bhan, 2019; Simone and Pieterse, 2018). For the purposes of this paper, we utilize the final definition, operationalizing it by considering low- and middle-income countries as the Global South and high-income countries as the Global North, though some of the scholarship we draw on uses a world-regional definition. As with any definitions of South and North, both containers are expansive and internally diverse, along social, cultural, political, historical, and climatic lines.

In examining urbanization patterns in the Global South, our reference period is the last 70 years. Global South urban transitions have taken place mostly since 1950, or even somewhat later, whereas they took place in most of the Global North in the 19th century. Therefore, in order to compare urbanization in the North and South, we will have to toggle between now-now and then-now comparisons. The former are necessary means for assessing how urban theories might apply across both geographies in contemporary times. Yet in evaluating the robustness of theories of urban genesis, we must compare the previous period in the Global North to the recent period in the Global South. In what follows, this is not an inconsistency, but a methodological choice, and by making it we are not implying that Global South urbanization is belatedly following the trajectory of the Global North; in

---

2 The urban transition is commonly defined as the period when a country becomes majority urban. Throughout the world, urbanization is a very old process, but the first urban transitions occurred in Europe during the industrial-manufacturing revolutions of the 19th century.
fact, our analysis ultimately supports claims that urbanization in the South is following pathways distinct in some ways from those experienced in the North (Edensor and Jayne, 2012; Robinson, 2006). But it also shows that there are some similar forces at work in the two contexts.

This paper cannot review all scholarship on the urban South. Much contemporary research on the urban South is ethnographic, illuminating the texture of everyday life and the interactions among culture, social relations, and institutions that produce particular lived experiences (e.g. De Boeck and Plissart, 2014; Ghertner, 2015; Simone, 2010). We note linkages between these experiential qualities of urban South environments and the variables we have chosen to analyze, but we will not attempt to assimilate all this knowledge into the theoretical frameworks we use, nor would we suggest that theory is capable of doing so in either South or North. Moreover, as Lawhon and Truelove (2020) have argued, even the scholarship they term the “Southern urban critique” contains a range of perspectives, some of which emphasize empirical differences between South and North and others of which focus on the politics of knowledge production. Our entry point is primarily the former concern. However, by showing that an analysis of fundamental or general processes of urbanization can sharpen our understanding of the specificities of urban transitions in the South, we hope to also contribute to a broader conversation about knowledge production. Our paper is a reminder that social scientists can produce meaningful knowledge about difference by starting with the identification of processes, relationships, and dynamics that can be found in many different historical and geographical contexts.
2. Global South urbanization through the lens of global urban theories

In this section, we concentrate on four variables—demography, economic development, migration, and the production of the built environment—that are linked to urbanization in both the South and North. As noted above, we expect all of them to be generally relevant but to take the form of a distribution, thus incorporating difference and variation.

2.1 Demography and urbanization

Urban population growth in today’s South is without historical precedent. Between 1875 and 1900, the period of swiftest urbanization in the Global North, its urban population grew by about 100 percent; in the equivalent period in the South, between 1950 and 1975, its urban population grew by 188 percent (Williamson, 1988). Comparing major cities in their respective periods of rapid growth, the difference in absolute numbers becomes especially stark. For example, while Chicago and Mexico City experienced similar rates of growth in their periods of rapid expansion, 1875-1900 and 1950-1975, respectively, the former added 1.3 million people while the latter added 8.1 million (Brockerhoff and Brennan, 1998). These figures are not mere administrative artefacts, as studies of global urbanization patterns that use harmonized definitions of “urban” find similarly rapid rates of urbanization in the South (OECD and European Commission, 2020).

---

3 Neither the South nor North is homogeneous. We focus on generalizations that apply across large swaths of the South, knowing they do not apply equally everywhere within the South.

4 There is less consensus about the pace of urban transition—how quickly countries in the South have shifted from low to high ratios of urban to total population. Some argue that rates of growth in this ratio are similar when comparing historical periods of most rapid urbanization (Brockerhoff & Brennan, 1998), but others find that the South has undergone its urban transition more swiftly (Dyson, 2011).
The unprecedented scale and pace of urban population growth in the South is largely a function of overall demographic growth in Global South countries—not faster rates of rural-urban migration (Table A1 in online appendix). During much of the North’s urban transition in the 19th century, cities were death traps where mortality outpaced fertility (Davis, 1973); their population increase was mostly due to replenishment from in-migration from the countryside (or, in the case of former settler colonies like the United States, from other countries). Only 25 percent of urban population growth in Western Europe and North America from 1830-1920 was due to natural population growth in cities (Jedwab et al., 2017).

In contrast, natural increase in the South has comprised about 60 percent of urban population growth since as far back as 1960, and this figure continues to rise in most countries (Chen et al., 1998). Urban transitions in the South have taken place almost entirely in the presence of modern medicine and urban sanitation, reducing endemic early life mortality and generating longer life spans. Most (though not all) countries have escaped the Malthusian food scarcity trap.5

This different demographic context of urbanization is expressed in two ways. First, high rates of natural population growth in urban areas lead to rapidly expanding “mushroom cities” (Jedwab et al., 2017). Second, high rates of natural increase in the countryside can generate urban population densities in agrarian regions, as in much of contemporary South and Southeast Asia and West Africa—what Qadeer (2000) calls the “ruralopolis.” When these population dynamics combine with a local economic transition away from agriculture, the result is the urbanization of once-rural settlements, sometimes referred to as in situ urbanization. The demographic conditions of 19th-century Europe, let

5 There is no indication that even the worst public health emergencies of the 21st century—such as the Covid-19 pandemic—are capable of fundamentally altering the demographic trajectories of cities.
alone North America, would not have enabled this type of urbanization. Thus, earlier
declarations based on the North’s experience, to the effect that “depopulation of rural areas”
(Davis and Golden, 1954: 10) constitutes a universal feature of the urban transition, no
longer hold up in light of the Global South experience.

Nonetheless, this finding does not negate a powerful general fact of urban
transitions: they are driven by demographic transitions. The demographic transition, which
Kirk (1996) calls “one of the best-documented generalizations in the social sciences” (361),
involves a period of rapid population growth that societies experience when mortality
begins falling even as fertility remains high.6 Even in Europe, what eventually enabled the
creation of majority-urban societies was the reduction of urban mortality rates, such that
rural-urban migration was not merely replacing urban dwellers dying from disease but
actually expanding cities’ populations. Europe’s urbanization faced an impenetrable ceiling
before the onset of the demographic transition (Dyson, 2011). Demographic transitions, and
the massive population growth they entail, are similarly at the root of urban transitions in
the Global South today, even if their scale is greater and some of their empirical
expressions—“mushroom cities” and in situ urbanization—are historically specific. As Fox
(2012) explains: “Urbanization should be viewed as a global historical process propelled by
technological and institutional changes that alleviated the surplus and disease constraints
that limited urban population growth in the pre-industrial era” (303). Through a
comparative lens, one can see that the pace and scale of urban population growth in the
Global South is driven by swifter and greater rates of mortality reduction than during the

---

6 Note that early demographic transition theory made the error of assuming that fertility decline was the result of
industrialization and urbanization, whereas contemporary demographers largely agree that mortality decline is
the remote cause of fertility decline. See Dyson (2011) for more detail.
19th century in Europe, but not a fundamentally different structural relationship between demography and the urban transition. This general relationship is still capable of embracing a wide empirical elasticity, shaped by national and local policies, technology, climatic conditions, and a host of other interacting variables.

### 2.2 Economic development and urbanization

A staple of urban theory is that urbanization is positively related to overall economic development, meaning growth in per capita income. This relationship remains strong across the world today (Bloom et al., 2008), and there is general agreement that the positive association between the two is related to economies of agglomeration (Jacobs, 1969; Marshall, 1890). People and firms are more productive when located in close physical proximity to one another, due to the mechanisms of sharing, matching and learning (Duranton & Puga 2003). Because of these mechanisms, urban areas are productivity peaks within their national economies. Though congestion costs can detract from such productivity gains, there has yet to be shown a case anywhere in the world where these congestion costs outweigh the benefits of agglomeration, such that average urban incomes are lower than rural incomes.

As in the Global North, cities in the South generally have the highest per capita incomes and wealth levels among places in their respective national economies, mostly because they concentrate the most productive economic activities. Agglomeration benefits,

---

7 Though the direction of causality is debated (Bloom et al., 2008).
8 This argument does not imply that urbanization alone is a sufficient condition for economic growth; because growth rates are affected by many factors other than urbanization, countries experiencing low or even negative growth can continue to urbanize (Fay and Opal, 1999). Although even in these cases, urban incomes remain higher than rural incomes.
despite claims to the contrary (Parnell and Pieterse, 2016: 236), are shown to be even larger in the Global South than Global North (Jones et al., 2020). Chauvin et al. (2017) demonstrate that there is also a size premium on productivity among cities in the South, with the biggest cities gaining the most. Education and skills contribute to productivity and wages regardless of the size and density of a settlement, but when they are concentrated in large, dense clusters of economic activity, there is a distinctive urban premium on productivity. This premium is even higher in major Global South countries today such as China, Brazil and India, than it is in the Global North (Chauvin et al., 2017). The aggregate benefits of agglomeration, therefore, are not specific to one world region or time period but rather appear to be features of urbanized economies everywhere.

Despite this association between urbanization and economic growth, it turns out that urbanization today occurs globally at lower levels of income than urbanization in the past. Jedwab & Vollrath (2015) find that, at every level of per capita income, the share of a country’s population living in urban settlements is 25 to 30 percentage points higher today than in 1500. In other words, if we compare a poor country today to one with similar income levels five centuries ago, the contemporary country would be far more urban than its historical counterpart. To understand this phenomenon and its consequences for the contemporary Global South, we must combine historical analysis with theoretical approaches from development studies and economic geography; we turn to this task below.

2.2.1 Urban economies in the global division of labor

While some exceptions can be identified—such as 19th-century Dublin (Larkin, 2000) or Naples (Davis, 2007)—a general fact of the Global North’s urban transitions is that they were shaped significantly by the growth of manufacturing. Factories provided urban
employment for rural migrants with limited formal education. In the First Industrial Revolution (textile-based) from about 1800-1870, this often involved the “immiseration” decried by mid-19th century observers such as Marx, Dickens and Zola, because the factory system largely replaced the existing artisan class. But with the advent of the skill-biased Second Industrial Revolution around 1880 (mechanical-, electrical-, and fossil-fuel-based)—in association with the growth of collective bargaining—urban manufacturing employment began providing pathways into the middle class. Though this relationship has since weakened, it was particularly important from 1880 to 1970 in the North.

A key reason why manufacturing in the North ultimately led to the creation of large urban middle classes is because cities of Western Europe and North America hosted the most advanced economic activities of their time. Because the knowledge and physical capital basis for their productivity was scarce, these cities’ economies enjoyed technological rents and favorable terms of trade. These rents were enhanced by imperialism, which used low-cost (including enslaved) labor to furnish cheap inputs to manufacturers and granted monopolies to European firms supplying colonial markets. Although the internal income distributions of these Global North cities were highly unequal, labor movements and public policies ultimately channeled economic rents toward substantial wage growth for workers, such that Northern industrial cities became, on average, the richest places in the world.⁹

In contrast to manufacturing cities of the historical North, most contemporary Global South cities that specialize in manufactures or tradable services produce goods that are already technologically mature. Their firms enjoy no extraordinary economic rents because their outputs can be produced in many places, leading to deteriorating terms of trade—often

⁹ With the notable addition of Buenos Aires in the early 20th century.
called a “race to the bottom.” To complicate matters, technological change and automation are now sharply reducing manufacturing employment-to-output ratios all over the world. This trend is leading Southern economies to deindustrialize at lower income levels than Northern economies did some five decades ago (Rodrik, 2016). Though some parts of the Global South, as in China, are making the transition to innovation-based industrialization, many others are not; thus, the pathway to generating industrial employment with increasing wages in many cities of the South is full of roadblocks. Similar phenomena can be observed in some Global North deindustrializing city-regions, such as the Rust Belts of the United States and Western Europe (Autor et al., 2016). But the downward pressure on wages and working conditions that these cities are experiencing starts from a point of relative wealth, and at least to some degree is compensated for by institutions of redistribution in national economies that remain wealthy by global standards.

Despite these challenges, Global South cities that host specialized manufacturing and tradable service sectors are in some ways the most opportunity-rich urban labor markets of the South, and they are the exception rather than the rule. Many cities in the Global South lack robust tradable sectors and face even steeper challenges in relation to the world economy. Especially in Sub-Saharan Africa and South Asia, urbanization with low levels of industrialization is widespread (Gollin et al., 2016; Kessides, 2007). Most of the informal non-farm work available in the South—comprising about half of all urban employment, compared to 17 percent in the North (see Table A2 in online appendix)—is in non-tradable sectors such as retail, hospitality and domestic work (Vanek et al., 2014). Upward mobility for workers in these sectors is even more constrained than in low-wage tradable industries, partly because they are likely to be self-employed (Munck, 2013). Nearly half of non-farm workers in the poorest Global South countries are own-account (neither employee nor
employer)—the figure is about 8 percent in the North (Gindling and Newhouse, 2014)—which limits the scope for collective bargaining. Even in cases where a Global South economy specializes in exports of a natural resource whose scarcity offers a comparative advantage, employment is nevertheless concentrated in non-tradable informal services, since 21st-century technology renders natural resource sectors capital- rather than labor-intensive (Gollin et al., 2016).

2.2.2 Trade costs and specialization

Economic specialization is the key link between agglomeration and income growth. Urban size and density enable sectors to deepen the division of labor, which raises productivity within and between firms (Spence et al., 2008). This process is closely linked to transport costs. As inter-urban transport costs decline, specialization can also take the form of an increasingly complex division of labor not only within but across cities (Henderson, 2005), a relationship formalized by the New Economic Geography (Fujita et al., 1999; Krugman, 1991). This is why urbanization and trade go hand-in-hand, within cities, between them, and across international borders (Spence et al., 2008; Storper, 2013).

Countries of the Global South commonly face higher internal trade costs than those of the North, leading to lower levels of urban economic specialization within national economies and lower overall productivity in cities (World Bank, 2009). Yet from a historical perspective, even in many Global South countries that host poor infrastructure relative to contemporary world standards, trade costs are much lower than the best of what was available in the 19th century North, due to motorized vehicles and telecommunications. The global reduction in trade costs helps to explain Jedwab & Vollrath’s (2015) finding that
urbanization levels are much higher today than in the past at every level of per capita income.

However, there is a cruel irony here. Lower global trade costs also trap many Global South cities at lower ranks in the global division of labor we refer to above because they render economies more open to global import competition. Venables (2017) finds that these pressures make it difficult for certain cities to successfully specialize through trade, as lower transport costs remove their historical dominance in serving regional or national hinterlands. As a concrete example, consider how challenging it is for a contemporary city in Tanzania or Nepal to compete with China or Vietnam in apparel production, even in selling to its own national market. Similarly, Glaeser (2014) argues that in closed economies, urbanization increases with rising agricultural productivity, but that in the open economies of the globalization era, this relationship is eroded or even inverted. Because of low-cost global trade, countries no longer need a large agricultural surplus to urbanize. Global North cities, during their periods of rapid growth, also participated in a global economy, but their major expansion occurred in an era when trade costs rendered national economies less open to long-distance competition, especially for tradable manufactures. When Global South cities do find a foothold for specialization in the global economy, the benefits can be circumscribed by the forces we identified above.\textsuperscript{10}

The demographic trends we refer to in section 2.1 compound the effects of these economic geography factors. Structural challenges to creating high-quality employment in the Global South come at a time when the working-age population in most Southern countries is expanding. Because the North’s urban transition involved the depopulation of

\textsuperscript{10} In some parts of the world, internal trade costs were lowered by colonization, while in others the damage wrought by colonialism is partially responsible for today’s high internal trade costs.
the countryside, the bargaining power of urban workers was eventually boosted as they faced decreasing wage competition from potential rural migrants. In Europe, this effect was augmented by out-migration to the Americas (Williamson, 1998). In large parts of today’s South, rural regions remain labor-abundant, exerting downward pressure on urban wages, and any putative relief from international out-migration is limited by tight regulation of international migration. Even in today’s “left-behind” deindustrializing regions of the North, shrinking employment coincides with contracting, not booming, population.

These factors—a position in the global ladder of skills and technology that leads to unfavorable terms of trade, technological advancements resulting in lower demand for labor, and dramatic declines in trade costs that disadvantage unspecialized urban economies—all point to a different economic context of urbanization in the South today, as compared to the North, past or present. Yet these differences are revealed rather than obscured through analysis of the relationships that underpin a unified theory of urban economic development, consisting of interactions among agglomeration, trade costs, specialization, technology, demography, and overall development. Moreover, none of these differences contradicts the general relationship between urbanization and economic growth that persists across South and North.

### 2.3 Urbanization and migration

Migration is also fundamental to the urbanization process around the world, today and in the past, both during a country’s urban transition and as its system of cities expands and is restructured. Migration is shaped by “pull” factors that draw people toward cities and “push” factors that compel them (sometimes forcefully) to leave their places of origin.
Two theoretical frameworks enable a global comparative analysis of the relationship between migration and urbanization. First, Zelinsky’s (1971) concept of the mobility transition, though it suffers from some of the overreach common to modernist stage-of-development theories (Cooke et al., 2018), contains insights that prove useful in examining mobility and urbanization. Zelinsky posited that as societies go through a demographic transition, they gradually move from low to high levels of permanent internal migration. This squares with evidence showing that rates of permanent residential relocation in the Global South, where population booms are still unfolding, are lower than they are in the Global North (Bell et al., 2015). He also argued that early in the demographic transition, rural-urban migration plays a larger role in urban growth, but later on, cities’ populations expand primarily by natural increase and inter-urban migration. Here too, the claim is generally accurate; countries in the Global South have seen the contribution of rural-urban migration to urban population growth gradually fall in the course of their urban transitions, and rates of urban-urban migration are now generally on the rise (Deshingkar and Grimm, 2005; Zhang et al., 2020). And as Zelinsky’s theory predicts, the Global North has higher levels of ongoing city-to-city sorting of individuals and households.

Zelinsky’s framework is nonetheless unable to provide a full explanation for the wide variation in internal migration patterns that we observe across the world. One key difference between South and North is that the shift to natural increase-driven urban population growth has occurred at lower levels of development in the South than in the North, and with large shares of the population still living in rural areas. In much of the South, rural populations have continued to expand alongside urban populations rather than diminishing through an exodus to cities, and large rural-rural migration flows have developed (Kelly, 2011; Potts, 2013). Indeed, many scholars of migration and development
are grappling with the question of why rural dwellers are not migrating in greater numbers to cities, given large rural-urban wealth gaps (Kone et al., 2018; Munshi and Rosenzweig, 2016; Potts, 2016).

Zelinsky saw high levels of inter-city circulation as a feature of high-income countries—defined by leisure trips, business travel, and long-distance commuting. Yet “circulatory urbanism,” as termed by Srivastava & Echanove (2013), is a common feature of Global South countries where poverty and informality remain prevalent—a product of multi-sited livelihood strategies compelled by the economic conditions we describe above. In parts of South Asia, Southeast Asia and Sub-Saharan Africa, temporary and circular migration are, or have been at various moments during the urban transition, far more common than permanent rural-urban migration (Beguy et al., 2010; Blumenstock, 2012; Deshingkar, 2006; Hugo, 1982). In India, for example, some scholars estimate that about 100 million people are circular migrants (Deshingkar and Akter, 2009). Circular migration also occurred in cities of the 19th-century Global North (Tilly, 1976), but given dramatic improvements in transportation and communication technologies, there is likely an order-of-magnitude difference in the scale and frequency of these back-and-forth movements in today’s South (Sheppard, 2014).

Zelinsky himself acknowledged limitations in his mobility transition theory (Woods et al., 1993), and we can look beyond Zelinsky to draw on another general theory of internal migration and urbanization. This framework relates underlying preferences for income, quality of life, proximity to non-pecuniary advantages (such as family), and other kinds of amenities and dis-amenities to the ongoing sorting of people, households, and economic
activity among and within regions (Glaeser, 2008). Migration decisions are the result of trade-offs people make among these preferences, which are also subject to financial constraints, laws, and cultural norms. The choice sets of this spatial arbitraging process are shaken up by shocks from technology and job location, climate change, crime, housing prices, conflict, and a host of other forces.

This framework directs our attention to the locational choice sets that individuals and households face when they consider whether to move to or between cities. A key difference between South and North is the overall structure of urban systems and the migration choice sets they offer to households and firms. Urban systems in the South tend to host large concentrations of people in either major cities or small cities and towns, with few prosperous middle-size cities (OECD and European Commission, 2020; Randolph and Deuskar, 2020; World Bank, 2009). This difference in urban system structure reflects a general relationship between income level and the structure of urban systems. As countries undergo economic development, they tend to generate a wider availability of employment, income and amenity opportunities (Williamson, 1965). The polarized rank-size structure of many Global South urban systems also reflects the historically specific economic development headwinds that we outlined above. In an era of globalization and automation, Global South countries confront steep challenges in creating the conditions for productive employment across a wide range of specialized cities—leaving individuals and households with a restricted choice set of migration destinations.

---

11 Some urban economists view the outcomes of such arbitraging as leading to a specific version of spatial equilibrium, in the form of an urban system featuring equalization of “real utility” among regions. This is a strongly normative framework, and we consider that version of the spatial equilibrium framework to be highly problematic, in analytical, empirical and historical terms (Storper, 2013).
This restricted choice set means that highly skilled workers in the South are even more likely than their Northern counterparts to be drawn to the biggest cities (Dingel et al., 2019; Fu and Gabriel, 2012). This also reflects common global trends of increasing urban size premiums on the wages of high-skilled workers (Autor, 2019; Chauvin et al., 2017) and the increasing urban specialization of the biggest metropolitan areas across the world (De La Roca and Puga, 2017; Diamond, 2016; Storper, 2013). But much of the workforce in the Global South struggles to benefit from the productivity of large metropolitan economies given their narrow pathways to economic mobility for those without formal skills and education. Many individuals in the South thus choose to remain in place, hence the persistent population growth in much of the rural South. Many others straddle multiple locations by moving seasonally or commuting long distances between their rural origins and other labor markets, both rural and urban, that offer mostly informal, insecure employment opportunities—helping to explain patterns of circulation.

Whether or not internal migration and spatial sorting patterns in the South and North converge over time depends on how urban systems in both contexts evolve. Given the structural factors we identified in the previous section, the South may not create a wide range of medium-sized specialized cities offering productive employment opportunities, just as the North may see a continued hollowing out of the middle of its urban systems amid deindustrialization. Climate change will also influence the migration landscape of countries in both South and North (Call and Gray, 2020; Henderson et al., 2017). Climate shocks may further reduce the locational choices available to households, reshaping rural-urban, urban-urban or urban-rural migration flows within countries and between South and North.

The mobility transition and spatial sorting theories cannot explain everything about a phenomenon as heterogeneous as human mobility or the structure of an urban system. As
we have previously stated, theory, in its search for generalization, does not provide complete explanations for real-world complexity and variety. Instead, these general theories are useful as workhorse models. They build a foundation for further analysis, which would require rich and varied information on the historical and geographical specificities of migration across the Global South and North.

2.4 The production of the built environment

The construction of the built environment, its spatial form, and its functional organization are key to the urbanization process across time and space. Cities are vast concatenations of buildings and infrastructure. These take an enormous variety of forms and give rise to experientially diverse urban environments. Scholars of the Global South have developed a vast literature on the built environment, where we observe the most striking differences between Southern and Northern cities. The production of the built environment in Southern cities often follows a sequence that represents the “reverse order” of physical urbanization processes in the North (Greene & Rojas, 2008). Many households in the Global South lack access to a formal mortgage market and therefore move into their homes when the structures host only the most basic features, upgrading them as savings or other small, informal loans become available (Amoako and Frimpong Boamah, 2017; SPARC, 2013). This method of incremental or “auto-construction,” as termed by Caldeira (2017), is “the process used by much of the low and moderate-income majority of most developing countries to acquire shelter” (Ferguson & Smets, 2010, p. 1).12

---

12 Auto-construction can be found commonly in a few areas of the Global North, such as southern Italy or Greece; but on average, it is more prevalent in the South than the North. Auto-construction was, however, prevalent in the past in much of the Global North.
With incremental construction typically occurring outside the direct scrutiny of formal planning or financial institutions, Global South cities see limited application of official building codes and land-use regulations (Arku et al., 2016; Wang et al., 2009). This mismatch between the rules prescribed by official planning and regulatory authorities and on-ground realities applies to neighborhood-level zoning as well. In many neighborhoods of the urban South, residential, commercial, and industrial activities are intermingled in spite of zoning laws that are generally unsupportive of mixed-use development (Baffour Awuah and Hammond, 2014; Liu and Liang, 1997; Sharma, 2000). As emphasized by Roy and AlSayyad (2004), flouting of official building codes and land-use regulation is not simply a survival strategy of the poor in the urban South, but is practiced across social classes and has become institutionalized in some contexts, in the sense that it is an expected part of the process. Even areas of the South that appear formally regulated often have widespread informality, such as central Shanghai’s informal market for crowded bed space rentals (Harten et al., 2020).

In this context, property titling systems have limited reach, and there are multiple, overlapping and often contradictory systems of land management (Buckley and Kalarickal, 2005; Mercer, 2017). As a result, residents of Global South cities face widespread tenure insecurity (De Soto, 1996), and many engage in an ongoing struggle for access to basic services. With formal service delivery networks failing to reach those with incomplete legal claims to their homes (Heller et al., 2015; Watson, 2009), residents are obliged to develop alternative solutions for needs like water, electricity and waste disposal (Schindler, 2017). The resulting improvisations become infrastructures that either parallel or intersect with formal networks, evolving as part of the built fabric of Global South cities (Bhan, 2019; Silver, 2014).
Land management and service delivery become formalized in some parts of the Global South in reverse order from the North. If they do not end in demolition or eviction, tenuous land claims and improvised service delivery may be the first stage in a multi-step process that proceeds to some form of official recognition, and then eventually reaches post hoc formalization, where rights to services and land are recognized. The liminal stage(s) usually involve(s) some kind of “permission”—which may itself rest on tenuous legal claims—that enables partial extension of services or tenure security without providing full rights (Criqui, 2016; Hylton and Charles, 2018). These trajectories from informal to formal contrast with the current process of urban development in the North, which usually begins rather than ends with land titling, permitting, and the provision of public services.

Certainly, exceptions can be identified in the North (and conversely in the South); auxiliary dwelling units in California began as informal housing and were later authorized post hoc (Mukhija, 2014), but they are rare in the Global North and common in the Global South.

These differences in how the built environment is produced can foster distinct rhythms of public life in Southern cities. Mixed-use informal neighborhoods, where private space is limited, family units are large, and work and residential spaces blur into one another, can generate more active streets. Kim (2015) records the intensive, creative and dynamic use of sidewalk space in Ho Chi Minh City, in contrast to most Global North cities where sidewalks are used less often, less intensively, and for fewer purposes. Akbar et al. (2018) use traffic data to show that Indian cities do not have two distinctive congestion peaks (a morning and evening rush hour), which they attribute to “the roadway [being] used for multiple purposes from late morning until well into the evening” (19). These findings echo Parnell and Oldfield’s (2014) suggestion that there is a different “human energy” (1) in cities of the South; Sheppard’s (2014) argument that streets of the urban South
are “seemingly chaotic from the perspective of northern cities” (148); and Simone’s (2010) description of a “a new urban sociality” developed “under dire conditions” (315).

The starting point for mainstream urban economics is that land is fundamentally different from other inputs and outputs of the economy, and this can be useful in understanding the processes described above. While markets for most goods and services adjust to increases in demand through increases in supply, land is immobile and hence positionally supply-inelastic (Ricardo, 1821). In the process of urbanization, density and proximity are the material form through which humans realize the benefits of agglomeration; land thus becomes especially valuable, but also contested and rivalrous, especially in central locations (Alonso, 1960). This results in a specifically urban political economy based on land use, where landowners attempt to shape and reshape land use for profit, residents attempt to improve or defend their quality of life, while other users groups pursue specific land use objectives (Logan and Molotch, 1987).

Scott (1980) integrated the economic and political dimensions of land in his theory of the “urban land nexus,” which maintains that there are numerous positive and negative externalities of proximity in cities that are not found in less dense land-use patterns, leading to social and political processes of conflict and accommodation (Scott & Storper, 2015). These negotiations generally deal with how and where the built environment is produced, the functional and spatial organization of the city, and specifically the land uses from which residents, or specific groups of residents, seek to separate themselves. Underlying motives may be socio-economic—for example, the rich desiring to separate themselves from the poor or from pollution—or functional—for example, separating a retail shopping district from a heavy manufacturing area. But these forces of separation or segregation are everywhere in tension with the desire for density and proximity, which allow cities to be productive and
interactive (across activities and groups) in the first place. Archaeologists hold that the "spatial division of cities in districts and neighborhoods is one of the few universals of urban life from the earliest cities to the present" (Smith, 2010, 137).

The urban land nexus provides theoretical scaffolding for analyzing the phenomena we have outlined in this section. The process by which an informal settlement is incrementally constructed and formalized involves a negotiation over urban land and competing claims to valuable space. One factor shaping this political struggle is the paradoxical position of the middle and upper classes in cities, who both depend on the proximity of labor and goods provided by the working class and seek to separate themselves from the poor. Elites in the Global South use various strategies of "partial recognition" in informal settlements, as outlined above, or campaigns to "sanitize" the city (Ghertner, 2015; Kundu and Saraswati, 2012) to grapple with these contradictory impulses. In the North, high-income groups often use other strategies to keep the working poor both near and separate, such as redlining or exercising influence over the location of public housing. The urban land nexus interprets these practices as expressions of a uniquely urban tension, *i.e.* the need for proximity and the desire for separation—mediated by class and many other social cleavages, such as caste, religion, race or national origin.

The urban land nexus theory cannot fully explain the detailed texture of land use in cities of the South or North. The specific patterns of proximity and avoidance, and regulation and governance, of the urban land nexus unfold differently across cities, shaped by a wide range of values, norms, institutions, histories, and natural geographies. They also interact with other fundamental factors of urbanization we have identified; for example, Balakrishnan and Pani (2020) argue that, in low-wage democratic societies of the Global South, "the state politically manages the crisis of joblessness by building constituencies
around negotiated access to land and housing” (3). Differences in how the built environment is produced are experientially and materially striking, but beneath them are variations on a common theme of tension between the forces of separation and attraction, between centripetal and centrifugal social interactions.

3. Urban theory, contextual analysis, and rigorous comparative methods

The strength of a theory can be judged by the range of outcomes over which its tools contribute meaningfully to explanation. To highlight and explain variation across time and space, we must first disentangle what varies from what does not and explore the distributions and magnitudes of variations or differences. To use a metaphor, the laws of engineering that determine load-bearing capacity and stability for walls are universal. Yet walls can be constructed of many different materials, and their design is influenced by different climates, cultural practices, and histories. In the end, walls are much more than engineering, but they cannot be explained without the laws of engineering. Social science, including urban research, struggles to establish causal relationships and empirical regularities in the ways that natural and physical sciences do, for we do not have the same sample sizes, time dimensions, or ability to conduct controlled experiments. The words “universal” and “law” rarely apply in the social sciences, and we intentionally chose not to use them in referring to theories of urbanization. Nevertheless, the search for generalizable theoretical relationships in the social sciences, including in urban studies, remains necessary and worthwhile for comparative research.

As noted in section 1, some scholars of the urban South have argued that significant empirical differences in Southern urbanization necessitate different theoretical approaches
(Robinson, 2002; Sheppard et al., 2013). Many assign the distinctiveness of Global South cities to an enduring legacy of European colonialism and contemporary forms of economic imperialism—such that theories developed elsewhere will always face a “post-colonial conundrum” in the South (Sheppard et al., 2013: 895). Colonization has had dramatic impacts on societies, economies, and institutions in both South and North—effectively reversing global patterns of prosperity prior to 1500 (Acemoglu et al., 2002)—and, hence, has profoundly influenced cities and urbanization processes. In much of Asia and Sub-Saharan Africa, European powers reshaped economies around agricultural and mining capitalism, restructuring urban systems in the process. Globally oriented port cities, whose primary economic function in colonial times was to export raw commodities to the metropole (and later to import manufactured goods), grew most rapidly, while the development of regional, manufacturing-oriented cities, some of which had long histories of proto-industrialism, was undermined (King, 1990). In settler colonies—where the model of extraction involved the persecution, dispossession, massacre, and ghettoization of indigenous or first nations peoples, then the imposition of systems of slavery and indentured servitude using human labor from Africa and parts of Asia—precolonial societies were so dramatically disrupted that colonial powers almost, though not fully, built settlement systems from scratch.

Colonialism’s legacies are extremely durable and widespread, both in the formerly colonized parts of the world and in European countries. In addition to having much larger impacts on processes beyond urbanization, these legacies shape the empirical expression of the theoretical relationships we outlined in section 2. For example, because colonization powerfully shaped the global division of labor and enforced favorable terms of trade for Northern economies, it created long-term advantages for Global North cities and long-term constraints to high-wage employment and large urban middle classes in Global South cities.
Colonial systems of administration also imposed various forms of spatial segregation in cities that have endured as context-specific articulations of the urban land nexus (Baruah et al., 2017). The legacy of colonization, however, is larger than any specific theory; rather, it is a macro-historical current that impacts the empirical expression of the theoretical relationships we have considered. The distinction we drew above between theory and explanation is useful here; explanation is different from theorization in that it requires us to place general theoretical relationships in historical and spatial contexts, in all their richness (Braudel, 1992; Ndiaye, 2009). In this sense, explaining any contemporary city requires that we examine how it was influenced by past colonization, especially the pivotal world-historical epoch of European colonization.

Global urban theory and context-sensitivity are not opposed to one another; they are inherent to understanding cities as the ultimate “noisy” social science problem. To continue our metaphor, just as common laws of engineering do not compel uniformity in architecture, understanding fundamentals of urbanization should not lead to uniformity in urban policy. On this point we agree with some scholars of the urban South (Robinson, 2006). Even on their own, the four relationships we examine in this paper combine in different ways and orders of magnitude from city to city, and especially between Global North and Global South, to produce specific urbanization contexts. Additional contextual factors also shape the terrain of urban governance: national policies (e.g. fertility policies, portability of social protection), institutions (e.g. democratic versus authoritarian regimes), physical geography (size, climate, topography), and socio-cultural factors (linguistic-cultural variation and division).

This observation leads to a final extension of our argument, which is to call for more large-scale and rigorous comparative urban research. This would enable us to better
understand the extent to which differences across cities are attributable to variation in the relationships we discuss in this paper (for example, the relationship between demography and urbanization); to combined, joint or indirect effects of such variation; or to variation in heretofore undiscovered fundamentals of urbanization. More comparative urban research would not only expose additional differences between the South and North, beyond those covered in this paper; it would also underscore important differences within the South and enable a constant interrogation of “Global South” and “Global North” as geographic containers. In the end, these are not fixed categories but analytical tools.

The identification of similarity, variation, and difference, with many interactions, is methodologically challenging, but social science offers several examples of rigorous comparative methods that aim for reproducibility of results. These include formal synthetic control models (Abadie, 2020; Abadie & Gardeazabal, 2003; Card, 1990); comparative macro history (Braudel, 1992; Moore, 1966; Tilly, 1984); comparative political economy (Acemoglu and Robinson, 2002; 2012); and, in a more constructivist vein, actor-network theory (Latour, 2005) or its cousin, social network analysis (Granovetter, 2005; White, 2008). While these are different approaches, they share a strong relationship to theoretical social science, in the sense that they seek causal generalization around fundamental relationships but look for variation and difference through interactions and pathways. Pressing questions confront the field that require deploying such tools for large-scale comparisons and generalizations. With the acceleration of technological advancements amid the Fourth Industrial Revolution (Schwab, 2015), how will the role of cities in agglomeration and specialization change across different societies? When the global population peaks near the end of the 21st century, how will this impact the dynamism of cities, as the structure of urban systems comes to depend entirely on the reshuffling of populations through migration rather than natural population
growth? How will the new ways of constructing and fortifying cities, necessitated by global climate change, impact the way they look, feel and function? All large-scale comparative research is intellectually ambitious and costly in terms of effort and data, yet the ability to conduct such analyses in urban social science is improving thanks to new data sources that utilize harmonized approaches to measurement (Charles-Edwards et al., 2019; OECD and European Commission, 2020).

4. Conclusion

Our exploration in this paper points to the promise of unified but flexible theoretical frameworks to yield deep insights about fundamental similarities in urbanization processes and sharper understandings of differences generated by empirical variation in theoretical relationships and the interactions among them. The fact that theories developed by studying Northern urban transitions cannot explain all dimensions of urbanization in the South does not betray any specific limitation of “Northern theory;” rather, it reflects the limitations of theory in general, regardless of its origin or where it is applied.

Our argument aligns with a recent intervention in this journal by Scott (2021), who contends that “urban phenomena are susceptible to investigation at the highest levels of theoretical generality” (1). We have built on Scott’s claims by demonstrating how interactions among key theoretical relationships are generating fresh insights into the empirics of urbanization in the Global South. In this sense, we concur with another recent Debates article by Fox and Goodfellow (2021), who describe how the conditions of “late urbanization” create specific challenges and opportunities for countries in the Global South. We contend that understanding how these conditions combine to produce the urban environments of today’s Global South—and especially how those conditions and their
empirical outcomes contrast with those of the historical North—requires application of the
types of general theoretical frameworks we deploy in this article. Disengaging from these
theories because of their geographic and historical origin would involve significant loss of
knowledge and insight into urbanization processes.

References

Abadie A (2021) Using synthetic controls: feasibility, data requirements, and methodological

10.1257/0028280321455188.


Acemoglu D, Johnson S and Robinson JA (2002) Reversal of Fortune: Geography and
Institutions in the Making of the Modern World Income Distribution. Quarterly

IDEAS Working Paper Series: 64.


Amoako C and Frimpong Boamah E (2017) Build as you earn and learn: informal urbanism
and incremental housing financing in Kumasi, Ghana. Journal of Housing and the Built

regulations in Accra-Tema city-region, Ghana: exploring the reasons from the
perspective of multiple stakeholders. Planning Theory & Practice 17(3): 361–384. DOI:
10.1080/14649357.2016.1192216.

DOI: 10.1257/pandp.20191110.
32


Potts D (2013) Rural-urban and urban-rural migration flows as indicators of economic opportunity in Sub-Saharan Africa: what do the data tell us? Migrating out of Poverty Research Programme Consortium. Sussex, UK. Available at: https://assets.publishing.service.gov.uk/media/57a08a2a40f0b652dd0005c8/WP9.pdf.


