THE JOURNAL OF TRANSPORT AND LAND USE http://jtlu.org
VOL. 4 NO. 2 [Summer 2011] pp. 47–69 doi: 10.5198/jtlu.v4i2.197

JTLU

Does first last?

The existence and extent of first mover advantages on spatial networks

David Levinson
University of Minnesota[a]

Feng Xie
Metropolitan Washington Council of Governments[b]

Abstract: This paper examines the nature of first-mover advantages in the deployment of spatially differentiated surface transport networks. A number of factors explaining the existence of first-mover advantages have been identified in the literature; however, the questions of whether these factors exist in spatial networks, and of how they play out with true capital immobility have remained unanswered. By examining empirical examples of commuter rail and the Underground in London, first-mover advantage is observed and its sources explored. A model of network diffusion is then constructed to replicate the growth of surface transport networks, making it possible to analyze first-mover advantage in a controlled environment. Simulation experiments are conducted, and Spearman rank correlation tests reveal that first-mover advantages can exist in a surface transport network and can become increasingly prominent as the network expands. In addition, the analysis discloses that the extent of first-mover advantages may relate to the initial land use distribution and network redundancy. The sensitivity of simulation results to model parameters are also examined.

Keywords: First mover advantage; Transport; Land use; Network growth

1 Introduction

The notion of first mover advantage (FMA) probably derives from chess or other competitive board games, in which the player who makes the first move has an inherent advantage (Streeter 1946). Since the introduction of modern game theory in the 1940s (von Neumann and Morgenstern 1944), first-mover advantage has developed into a game theoretic notion that, in a sequential round of strategic moves, a player may earn a greater pay-off by acting first rather than by following others. Examples include the two-stage Stackelberg Game and the Cournot Game (Gibbons 1992).

Over the last few decades, research on first-mover advantage has gathered momentum as the concept has found applications in the fields of industrial marketing and management. Since the publication of the seminal paper by Lieberman and Montgomery (1988), who define first-mover advantages in terms of "the ability of pioneering firms to earn positive economic profits" in a competitive market, a broad literature has been dedicated to exploring the mechanisms that confer advantages on first-mover firms in specific market segments (Kerin *et al.* 1996; Makadok 1998; Mittal and Swami 2004; Rahman and Bhattacharyya 2003). Extending the work of Lieberman and Montgomery, Mueller (1997) identifies a number of sources for first-mover advantages, which he sees as related to both demand and supply. Demand-related inertial advantages include set-up and switching costs, network externalities (effects that one user of a good or service has on the value of that product to other people), and buyer inertia (consumers' resistance to changing their buying choices) due to habit formation or uncertainty over quality, while supply-related efficiency advantages include set-up and sunk costs, network externalities and economies, scale economies, and learning-by-doing cost reductions. Controversially, first movers may also experience disadvantages; for example, early entrants to a market may miss the best opportunities and acquire the wrong resources, the optimal course of action being obscured by technological and market uncertainties during the early stages of the market (Lieberman and Montgomery 1998).

This study aims to examine the existence and extent of first-mover advantages in the deployment of spatial transport networks. It should be noted that transport systems differ from board games or industrial markets in many aspects, and the

---

[a] dlevinson@umn.edu
[b] fxie@mwcog.org

first-mover theories mentioned above may not apply. For instance, in contrast to firms in a free market that are private and highly competitive in nature, transport infrastructure is largely public, so providers of transport infrastructure may be neither profit-pursuing nor in competition with each other. While highway and transit systems are provided with extensive public funds, there are also privately developed transport systems such as streetcars and interurbans (Diers and Isaacs 2006; Marlette 1959). That being said, the notion of first-mover advantage needs to be carefully redefined for the purpose of this study based on an in-depth examination of the distinctive characteristics of spatial transport systems.

Spatial heterogeneity is an important feature of surface transport networks. In contrast to systems that can be deployed universally, transport infrastructure must be deployed in some place first in a spatially differentiated environment. Thus, infrastructure deployed in a particular place may gain an advantage because it has acquired the best location, or because it was first in time, or both.

Another important feature of a transport network is capital immobility. Transport infrastructure embeds high set-up and sunk costs, which help an incumbent if they are well located by establishing spatial monopoly and chasing off rivals, but they also make it more difficult for the incumbent to move, as they are physically bound to the location in which they have sunk costs.

While spatial heterogeneity and capital immobility confer inherent locational advantage on first-deployed transport facilities, first-mover advantages may also arise from the establishment of standards or technology lock-in. The first standards to be adopted first acquire advantages as others seek compatibility in order to obtain access to the field where the standards are applied, and in turn help to lock in those standards. An example of technology lock-in is railroad track gauges, which are now standardized at four feet eight and one-half inches (1435 mm) in Britain and North America—the same as the first steam railway, and a mere half-inch wider than the typical pre-steam tracks in the mining districts near Newcastle. This first-mover advantage lasted despite some railways trying alternatives (e.g. the Great Western Railway was originally built at five feet six inches (1676 mm)), and the first gauge used on a network tends to be adopted by most subsequent lines (Puffert 2002). Alternative gauges would have accommodated wider, taller, and faster trains more easily, but could not be deployed economically because of network lock-in, including requirements to rebuild expensive sunk infrastructure like bridges and tunnels to accommodate the wider gauge. Similarly, some areas may get a technology before others and acquire advantages as the technology is diffused and adopted elsewhere. A theory of technology diffusion suggests that technologies are deployed in a pattern resembling an S-shaped curve (Kondratieff 1987). There is a long period of birthing, as the technology is researched and developed, followed by a growth phase as the technology is deployed, and a slower mature phase when the technology has occupied all available market niches. Nakicenovic (1998), by plotting a large number of curves for transportation systems, showed that S-curves fit the temporal realization of transportation networks very well. This study, however, does not consider standards or technology lock-in in its analysis.

In recent years, the emergence of "network science" has opened up new approaches to understanding the advantages of first movers in an evolutionary process of network growth. There is ample evidence that many networks, such as the World Wide Web and metabolic networks, exhibit a "scale-free" structure (Newman 2003) in which some nodes act as "highly connected hubs" that are more important than others. Barrat *et al.* (2004) further points out that nodes "entering the system at the early times have always the largest connectivities and strengths," as those with greater connectivity are more likely to attract links subsequently added to the network. Although surface transport networks are somewhat different from these scale-free networks due to spatial constraints, this concept of preferential attachment sheds light on how the established advantage of first-deployed facilities in a transport network could be reinforced when the network grows over time and space.

This study re-examines the question of first-mover advantages in the context of spatial transport networks, and asks if early arrival is the cause of higher connectivities and confers advantages to transport facilities, and if these advantages remain or change with the passage of time. In studying transport systems—which require many years to develop, are subject to both spatial and network constraints, and are shaped by interdependent economic and political initiatives in deployment decision-making processes—we are behooved to examine these questions using an approach different from the game-theoretical methods widely adopted in the industrial economics literature. To that end, we construct a model of network diffusion and analyze first-mover advantages in a controlled environment. The next section reviews network diffusion models in the literature. To illustrate the idea based on empirical observations, the subsequent section investigates the presence (or absence) of first-mover advantages in a particular transportation case: rail in London. Then a model of simplified network diffusion process is developed and measures of first-mover advantages proposed, followed by simulation experiments and a discussion of results. The concluding section highlights our findings and indicates future directions of this research.

2 Literature review

It has long been recognized that the development of transportation and land use is a coupled process in which each drives the other. Transport infrastructure is deployed to serve the demand of moving people to their desired land use activities, so diffusion of transport networks cannot be divorced from the evolution of underlying urban spaces. Urban development has been examined by a long line of studies. The pioneering work by von Thünen (1910) examined a monocentric city and predicted the land use distribution surrounding the city core. Central Place Theory, introduced by Christaller (1933), demonstrated the emergence of a hierarchy of central places serving surrounding markets at minimum transportation costs. New Economic Geography (NEG), exploring the spatial distribution of economic activity, presented a synthesis of theories to explain processes of concentration and deconcentration of industries and workers [1]. Based on the theoretic investigations, a series of land use models have been developed to forecast land use development while considering transportation as a critical contributing factor. For instance, Anas and Liu (2007) developed a dynamic general equilibrium model of a metropolitan economy and its land use and transportation. Anas and Arnott (1993) demonstrated the model in a prototype version of the Chicago Metropolitan Statistical Area (MSA). Levinson *et al.* (2007) modeled the co-development of land use and transportation as an autonomous process in which the relocation of activities and the expansion of roads are driven by the independent decisions of individual businesses, workers, and road agents. Many of these integrated land use models have been applied in urban planning studies and some developed into commercial packages (Examples include START (Bates *et al.* 1991), LILT (Mackett 1983), and URBANSIM (Alberti and Waddell 2000)). A comprehensive review of these models has been provided by Timmermans (2003) and Iacono *et al.* (2008). In most of these models, the evolution of urban space has been played out as the outcome of the location decisions made by residents and businesses, in which accessibility plays an essential role (Hansen 1959).

While the interaction between transportation and land use is recognized as playing an essential role in transport development, the complexity of the relationship is such that meaningful analysis requires intricate modeling processes. This has given rise to another line of studies focused on modeling the diffusion of spatial networks while either ignoring land use or treating it as exogenous; over the last half-century, these studies have produced a broad literature that ranges across geography, urban planning, engineering, and network science. These studies are documented in a comprehensive review by Xie and Levinson (2009d).

The earliest models of network growth date to the 1960s and 1970s, when geographers attempted to replicate the growth of transportation networks in terms of their structural changes. Taaffe *et al.* (1963) proposed a four-stage model to describe the sequential process of road network development. The Taaffe model was applied to the Atlantic seaboard of the United States (Pred 1966) and then to the South Island of New Zealand (Rimmer 1967). Lachene (1965) developed a staged model of network development on a hypothetical transport network. Starting with an isotropic network of dirt trails and a more or less uniform distribution of economic activity, the model builds a road network to link settlements. While some less-used trails are abandoned, some become paved roads and eventually connect into a superior network of railroads or highways.

Lacking a deep understanding of the underlying growth mechanisms of transport networks, geographers in the early days had to limit their modeling efforts to heuristics and intuition as they sought to replicate observed structural changes. It was not until the introduction of travel demand modeling for urban planning studies that scholars and practitioners were able to investigate the "optimal" designs of roadway spacing, capacity, and system configuration. Boyce (2007) provides a historic account of the early development of urban planning models. In recent years, solution algorithms to user equilibrium have been widely used to solve network design problems (NDPs) (LeBlanc 1975; Yang and Bell 1998). NDPs are typically formulated as bi-level frameworks in which the lower level seeks to establish demand-performance equilibrium while the upper level seeks to identify optimal investment decisions subject to budgetary and other restraints. Constrained by computational ability, the set of investment choices has to be limited to a small size.

Since the "new network science" came onto the scene in the 1990s, interest has emerged in investigating network growth based on the concepts of preferential attachment and self-organization deriving from natural science. Agent-based sim-

---

[1] New Economic Geography was pioneered by Krugman (1992). See Fujita and Krugman (2003) for an overview of this emerging field; see also Anas (2004) for a critique.

ulation has provided an effective tool to represent network growth as an integrated process of interdependent initiatives, and has seen widespread application in modeling the spatial expansion of transport networks. The active-walker model (AWM) proposed by Lam and Pochy (1993) describes the dynamics of a landscape in which walkers as moving agents change the landscape according to some rule and update the landscape at every time step. The active walker model was adopted by Helbing *et al.* (1997) to simulate the emergence of trails in urban green spaces shaped by pedestrian motion. Yamins *et al.* (2003) presented a simulation of road-growth dynamics on a land use lattice that generated global features (such as beltways and star patterns) observed in urban transportation infrastructure. Zhang and Levinson (2004) examined the growth of the road network in the Minneapolis-Saint Paul (USA) metropolitan area with autonomously operating links, "backcasting" (predicting based on historical data) road expansions twenty years from 1978, and comparing the predicted network in 1998 to the real one. Yerra and Levinson (2005) and Levinson and Yerra (2006), developing an agent-based model of network growth, demonstrated that a road network could spontaneously evolve into an organized hierarchical structure from either a random or a uniform state. Xie and Levinson (2007b) developed an evolutionary model of network degeneration based on the posited "weakest-link" heuristic, in which individual links are created as autonomous agents and the weakest link is closed at each time step in an iterative process. The model was employed to explore the decline and abandonment of Indiana Interurban network (Xie 2008). Corbett *et al.* (2008), on the other hand, developed a model of network expansion based on the "strongest-link" heuristic, in which new links are added one at a time according to the potential to improve accessibility, and employed the model to simulate the expansion of the network of enclosed, elevated walkways (skyways) in downtown Minneapolis. Both "weakest-link" and "strongest-link" heuristics originate from the "greedy algorithm" (Cormen *et al.* 1990), in which locally optimum choices are made in a discrete optimization process at each stage with the goal of finding the global optimum. In both cases, evidence has been found that even based on myopic and local optimum decisions, the models replicated well the course of link abandonment or addition when compared to historical observations.

3 Illustrative examples

Four illustrative examples illustrate qualitatively and quantitatively different aspects of the existence or absence of first-mover advantages. These are rail in London, the global aviation and maritime systems, and roads in the Minneapolis-Saint Paul (USA) region, known as the "Twin Cities."

3.1 Rail in London

The world's first steam railway, the Stockton and Darlington, was constructed in England in 1825, and the technology soon spread. The first railway reached Greater London in 1836; by 1868, the city's rail system had reached 50 percent of its ultimate extent (in terms of number of London area stations, excluding stations that were later closed), and by 1912, the London rail network had reached 90 percent of its ultimate exten. The first Underground railway was opened in London in 1863, connecting stations of the various surface railways. Half of the ultimate number of stations were open by 1912, and 90 percent by 1948. These data are illustrated in Figure 1.

Figure 1 shows two graphs, one for the National Rail (surface) lines and one for the London Underground. Each graph contains three lines. The first is the cumulative share of the number of stations opened by year. For the London Underground, 0 percent of stations were open in 1862, and 100 percent were open by 1999. The second is the cumulative share of number of current boardings and alightings at stations by the year they opened. So the number for a given year represent the current boardings and alightings for stations that were open by that year. The third indicates the share of the number of connections (in the case of National Rail) or the number of lines using that station (in the case of the London Underground) by year.

The graphs clearly show that the share of cumulative ridership is greater than the share of cumulative connections, which in turn is greater than the share of cumulative stations. In other words, stations constructed early in the development of the network have more connections than those constructed later, and still more riders than later stations.

This finding supports the hypothesis that a first-mover advantage exists in the development of the London rail network. The early stations were generally well placed in areas that, at the time, generated more traffic. While land use patterns and demand have shifted in London (Levinson 2008b), the underlying pattern was of early stations serving the then-dense core, while the core has remained a dense employment center. The early stations, those in the core, are also more likely to have multiple connections, but the additional connections do not, in themselves, explain the additional ridership observed at these stations; rather, we need to look for an explanation outside the network—at land use, and the mutual reinforcement between land use and the network (Levinson 2008a).

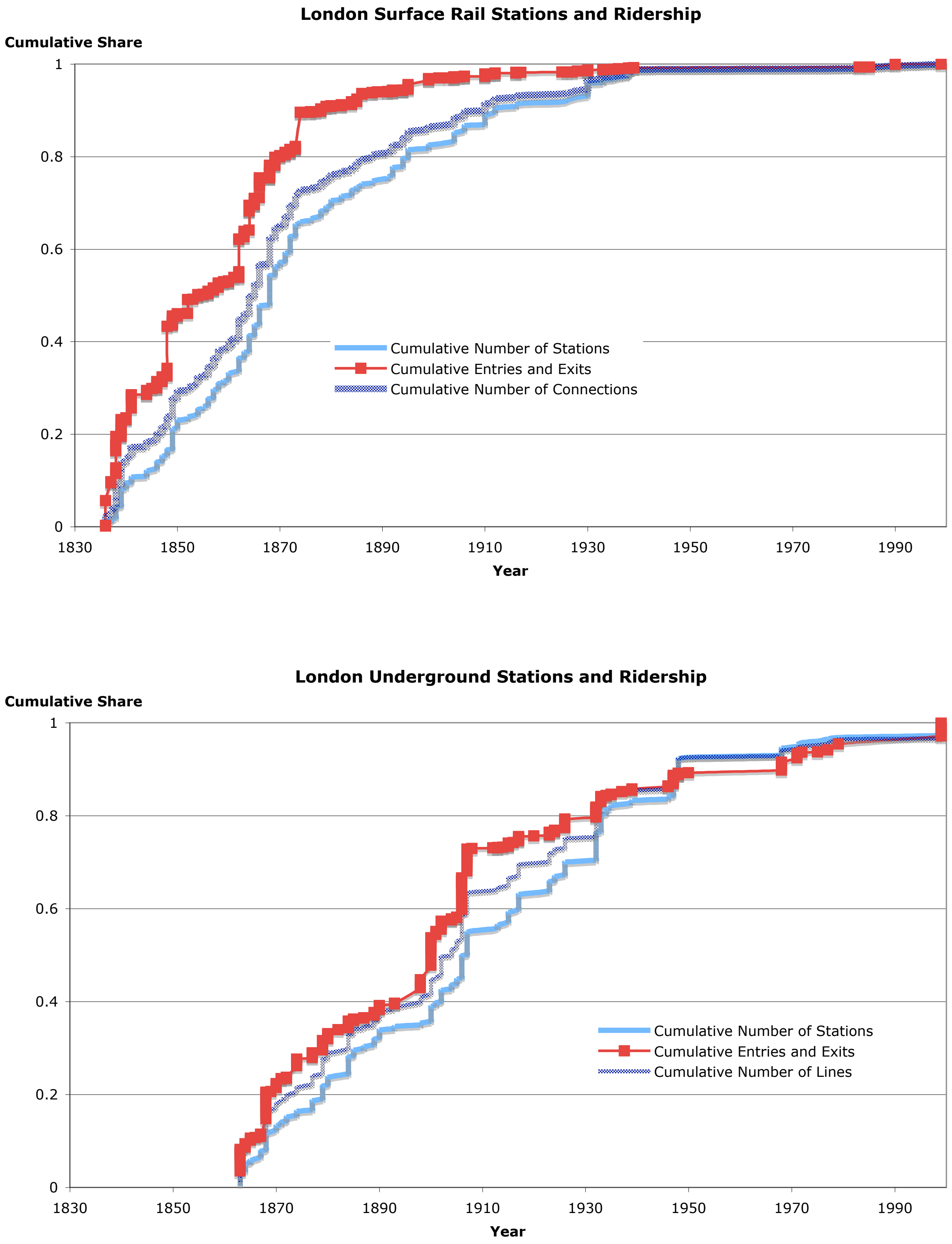

Figure 1: London rail cumulative stations and ridership. The first stations serve more riders than later stations. The first stations also have more connections than later stations, but not so much as to explain the additional ridership.

Examining the data[2] more rigorously in Table 1 suggests that the sources of the first-mover advantage are spatio-temporal location on the network and connectivity. Total station boardings and alightings on both the surface rail and Underground networks are positively related to number of connections and negatively related to travel time (in minutes) to Bank station (approximately at the center of the City of London) and statistically unrelated to year, after controlling for those two variables. Other variables controlled for were population and employment density (in thousands per km^2), which were insignificant, though notably highly correlated with station density; station density (in stations per km^2) (Underground station density was insignificant, surface rail station density was negative in both models, suggesting surface rail stations compete for customers while Underground stations complement, perhaps through higher densities); and location north of the Thames River (surface rail stations are somewhat less successful north of the Thames) and in the urban core (the London Boroughs of City of London, Westminster, Camden, Islington, Tower Hamlets, Kensington and Chelsea, and Southwark [3]. The entrepreneurs developing the rail system placed early stations well to take advantage of existing and prospective demand; the value of that placement remains today, a century after most stations opened.

3.2 Aviation

The global aviation system allows us to test first-mover advantage under particular conditions. There are a number of measures of airport size, including number of passengers, as shown in Table 2.

If there were a first-mover advantage, we would expect the oldest airports to be the largest. The data in Table 2 indicates otherwise: among the world's largest airports, there is no particular advantage to being a city with an earlier airport. Amsterdam-Schiphol, which opened in 1916, is ranked twelfth overall, while Dallas-Fort Worth (DFW), opened in 1973, is ranked sixth. One could argue that DFW is best considered a successor to an older airport: Love Field in Dallas opened in 1917 and began serving civilian flights in 1927, while Meacham Field in Fort Worth opened in 1925. However, those airports were different institutions located on different sites. Still, Dallas (along with Chicago) became a hub for American Airlines as early as 1930, by which time predecessor companies were already using that airport.

In contrast with first-mover advantages when comparing large airports, we can see significant persistence of hubs. Airlines that establish hub airports tend not to move them very often, and also crowd out other airlines seeking to establish hubs. For instance, Northwest Airlines was established in Minneapolis in 1926 and remains the dominant airline in that market eighty years later (under the name of its successor, Delta Airlines). American Airlines has remained similarly dominant in Dallas and Chicago. This persistence is not determinative; airlines with hubs do disappear (Eastern Airlines is a notable example), and do lose their hub advantage when faced with strong competitors. The case of US Airways versus discount carrier Southwest Airlines in Philadelphia is a telling example: David Siegel, the former CEO David Siegel of the then-entrenched US Airways said "They are coming to kill us," foretelling the loss of market share in Philadelphia (BTNews Online 2004), after Southwest had taken the Baltimore and West Coast markets from them. US Airways subsequently merged with America West, which took the US Airways name, but headquartered itself in Phoenix.

Within a metropolitan area, a first-mover advantage may be created when the first airport constructed captures the dominant share of (or even a monopoly over) locally generated traffic. However, that is difficult to test as so many cities have only one airport, and cities that once had more than one may

---

[2] Population data were obtained for the 33 current Administrative Districts (also called Boroughs, including the City of London and City of Westminster) of London. Density of population (and employment) were computed by dividing by the current area. This paper defines the surface rail system as all currently existing London-area heavy rail stations and lines that are not part of the 2006 Underground system, and the Underground stations are those that are part of the 2006 Underground system. Transport network data on the London Underground identifies each station on each line as a node, with X and Y coordinates, a date opened for a particular line. Very few stations were actually closed (as opposed to relocated), and these closures did not result in notable service reductions because the stations closed were those located too close to other stations. Underground stations that were opened and later closed were not considered as part of the analysis. Dates were obtained from (Rose 1983) and (Borley 1982). Small relocations of stations were ignored, as was the Circle Line, which shares platforms with the District, Metropolitan, and Hammersmith and City lines. When a new line was connected to a station, a new node was created for purposes of calculating station density, but each station (which may serve multiple lines) remained a single observation. The density of Underground stations was computed by dividing the number of Underground nodes (each station per line is treated as a distinct node, so a station serving three lines is counted as three nodes) at a given time by the current area. A similar procedure was used for surface rail stations.

[3] The core is defined as having a high degree of employment, areas where the ratio of persons working in the area to working-age residents exceeds one, (values in parentheses): City of London (55.74), Westminster (3.65), Camden (1.84), Islington (1.38), Tower Hamlets (1.16), Kensington and Chelsea (1.08), and Southwark (1.02). These areas are seven of the eight boroughs of London that have a ratio of jobs to working-age population greater than one; the other area is Hillingdon (1.16), which is located at the edge of the metropolitan area and is home to Heathrow Airport, and so is otherwise dissimilar from the core and is considered part of the periphery here) (Center for Economic and Social Inclusion 2006)

Table 1: London rail boardings and alightings regression model.

	Surface Rail Stations			Underground Rail Stations		
Independent Variables	Coefficient	T-Stat	P	Coefficient	T-Stat	P
Year	−3829	−0.4	0.69	25 881	1.55	0.12
Number of connections	1 850 620	10.45	0.00***	9 136 712	10.27	0.00***
Population density	160 836	0.84	0.40	446 271	1.46	0.15
Employment density	218 436	1.58	0.12	291 440	1.42	0.16
Underground station density	2 391 224	1.15	0.25	2 250 592	0.83	0.41
Surface rail station density	−7 369 758	−4.17	0.00***	−4 561 511	−2.45	0.02**
North of Thames [1,0]	−1 109 628	−1.79	0.08*	−1 197 065	−0.67	0.50
Core	−1 556 244	−1.26	0.21	112 884	0.06	0.95
Time to Bank station	−116 756	−3.8	0.00***	−211 722	−3.37	0.00***
Constant	9 077 731	0.51	0.61	−49 200 000	1.49	0.14
Adjusted R-squared	0.5339			0.4759		
N	308			257		

have relocated their traffic to a more suitable location. Alternatively, a second-mover advantage may accrue to a newer airport that is better suited to the changing local environment than older facilities located on small sites or facing high costs of rebuilding while remaining operational.

The surviving large airlines (network carriers) in the United States aviation system can trace their heritage to before the jet age, when air travel was uncommon and largely subsidized by airmail contracts with the Postal Service. Each airline has a distinct history, and consolidation has been common throughout the industry since its early days. American Airlines, for example, can trace its heritage to some 72 precursor companies. To illustrate with a simpler example, Northwest merged with Republic Airlines in 1986; Republic itself was the product of a 1979 merger between North Central Airlines (based in Minneapolis though founded in Wisconsin in 1939 and not moving to Minneapolis until 1952) and Southern Airways (founded 1949 in Augusta, Georgia) and a 1980 acquisition of Hughes Airwest, which was itself the product of a 1968 merger between Pacific Airlines (founded 1941 in California as Southwest Airways), Bonanza Air Lines (founded 1945 in Las Vegas) and West Coast Airlines (founded 1946 and based in Seattle). The Minneapolis hub was the first airport served by the original Northwest Airlines, while the original hubs or bases of predecessor airlines are no longer the dominant hubs of the current company (now incorporated into Delta Airlines).

Conversely, it is also possible to examine the first hub airports of the six airlines. The columns in Table 3 showing the first airmail and first passenger routes give insight into the first markets airlines occupied. These markets were allocated by the government (through either the Postal Service granting airmail contracts or the Civil Aeronautics Board allowing airlines to serve passenger markets). Those original markets are still dominated by the successor airline in five of six cases—the exception being Continental Airlines, which moved first to Denver, then pulled back from that city and then east to Houston (especially after its acquisition by Texas Air Corporation), a hub which it dominated.

Airlines, unlike airports, are largely composed of mobile capital. Despite the mobility of the main capital asset, airplanes, there is a tendency for airlines to persistently occupy hubs. This is demonstrated by the fact that all six major American air carriers are still serving at least one of their initial markets. The airline industry, because of its regulated nature through the 1970s, was a product of merger and consolidation as much as internal growth. The lock-in advantages, in addition to those noted by other authors above, include ownership or control of scarce gates at competitive airports, frequent-flyer loyalty programs that tie local residents to locally dominant carriers, and hubbing economies (a type of network effect) allowing hubs to provide frequent non-stop service to many cities.

3.3 Container Ports

Until containerization, longshoremen moved relatively small packages of goods on and off ships using cargo nets, grappling hooks, and brute force. The process of loading and unloading

Table 2: Airport Passengers 2006 (Top 30) by Opening Year.

Rank	Airport	Passengers	Year
1	Atlanta Hartsfield-Jackson	84 846 639	1925
2	Chicago O'Hare	76 248 911	1942
3	London Heathrow	67 530 223	1946
4	Tokyo Haneda	65 225 795	1931
5	Los Angeles Intl.	61 048 552	1929
6	Dallas-Fort Worth	60 079 107	1973
7	Paris C. de Gaulle	56 808 967	1972
8	Frankfurt	52 810 683	1936
9	Beijing Capital	48 501 102	1958
10	Denver	47 324 844	1989
11	Las Vegas McCarran	46 194 882	1942
12	Amsterdam Schiphol	46 088 221	1916
13	Madrid Barajas	45 500 469	1928
14	Hong Kong	44 020 000	1998
15	Bangkok Suvarnabhumi	42 799 532	2006
16	Washington, D.C. G. Bush	42 628 663	1969
17	New York John F. Kennedy	42 604 975	1948
18	Phoenix Sky Harbor	41 439 819	1935
19	Detroit – Wayne Co.	36 356 446	1930
20	Minneapolis-Saint Paul	35 633 020	1921
21	Newark Liberty	35 494 863	1928
22	Singapore Changi	35 033 083	1955
23	Orlando	34 818 264	1974
24	London Gatwick	34 172 489	1936
25	San Francisco	33 527 236	1927
26	Miami	32 533 974	1928
27	Tokyo Narita	31 824 411	1978
28	Philadelphia	31 766 537	1925
29	Toronto Pearson	30 972 566	1939
30	Jakarta Soekarno-Hatta	30 863 806	1984

Source: Airports Council International (2006)

might keep a ship in port for weeks. This "break-bulk" shipping was a major bottleneck in world commerce.

Malcolm McLean, a truck driver from North Carolina, conceived of loading trucks directly onto ships, without packing and unpacking, in effect using ships as transoceanic ferries. McLean realized that if the wheels were removed and the sides reinforced, trailers could be stacked. In April 1956, McLean's first container ship sailed from New York to Houston.

Containerization was essentially complete in 1971, when all containerizable cargo on the trans-Atlantic route was containerized (Rosenstein 2000). Yet the revolution continued as both the quantity of shipped freight and the size of the ships (and the ports required to accommodate them) grew.

The scaling made many older, smaller ports obsolete and created a new generation of superports that acted as hubs in a packet-based freight transportation system. Table 4 shows container port size in 1969, near the beginning of containerization. One notes, for instance, that Oakland had already beaten its competitor across the bay in San Francisco to containerization. Table 5 shows container port size in 2005, and a different picture emerges: only four of the top ten ports in 1969 (denoted in bold in both tables) remained in the top twenty, and only two in the top ten. Oakland, the second-largest container port in 1969, fell out of the top twenty as Los Angeles rose to take market share on the West Coast of the United States. The Australian ports of Sydney and Melbourne also fell off the list; Yokohama (Japan) was displaced by the slightly larger neighboring Port of Tokyo; Bremen (Germany) was replaced by Hamburg; and Felixstowe (southeast England) also fell off the list.

The new ports on the list are all from East or Southeast Asia with the exception of Dubai, which has emerged to fulfill a transshipment role for the Middle East.

What does this say about first-mover advantages? Ports are immobile capital, and while a port is certainly an important factor in a city's growth, it cannot alone determine that growth. As city-regions grow, and some specialize in producing or distributing tradable goods suitable for containerization, their ports will similarly grow. A port that grows early may retain some disproportionate advantage for a time while equilibrium is established; this advantage may carry over to other complementary aspects of manufacturing and trade, helping reinforce the port's position. The evidence, however, suggests that first-mover advantages are quite weak in this sector.

Table 3: United States network airline hub cities.

Airline	Year	Hub Cities	First mail service	First passenger service
American Airlines	1930	**Dallas**, Miami, San Juan, **Chicago**, St. Louis	St. Louis, Chicago	Dallas, Chicago, Boston
United Airlines	1926	Chicago, Denver, Washington (IAD), San Francisco, Los Angeles	Boise, Pasco	Chicago, Kansas City, Dallas
Delta Airlines	1924	Atlanta, Cincinnati, *Salt Lake City*, New York (JFK)	Fort Worth, Atlanta, Charleston	Dallas, Jackson
Continental	1934	Houston, *Newark*, Cleveland		El Paso, Las Vegas, Albuquerque, Santa Fe, Pueblo
Northwest Airlines	1926	**Minneapolis**, Detroit, Memphis, Tokyo, Amsterdam	Minneapolis, Chicago	Minneapolis, Chicago
US Airways	1939	*Charlotte*, Philadelphia, *Phoenix*, Las Vegas, **Pittsburgh**	Pittsburgh	Pittsburgh

Source: Airline websites.

Note: **bold** indicates original airport served by airline. Other hubs were often served by acquired companies; *italics* indicates original airport served by an acquired company.

Table 4: Container port size, 1969.

Rank	Port	Container Cargo (Metric tons)
1	**New York/New Jersey**	4 000 800
2	Oakland	3 001 000
3	**Rotterdam**	2 043 131
4	Sydney	1 589 000
5	**Los Angeles**	1 316 000
6	**Antwerp**	1 300 000
7	Yokohama	1 262 000
8	Melbourne	1 134 200
9	Felixstowe	925 000
10	Bremen/Bremerhaven	822 100

Source: Levinson (2006)

3.4 Twin Cities Roads

The Minnesota Department of Transportation (and predecessor organizations) have been building and maintaining roads in the Twin Cities (Minneapolis-Saint Paul) region since 1921. We have assembled a database of road projects by section, year built, and current utilization (measured as average daily traffic volume). The results of an analysis of this data,

Table 5: Container port size, 2005.

Rank	Port	TEUs (000s)
1	Singapore	23 200
2	Hong Kong	22 430
3	Shanghai	18 090
4	Shenzhen	16 200
5	Busan	11 840
6	Kaohsiung	9471
7	**Rotterdam**	9300
8	Hamburg	8086
9	Dubai	7619
10	**Los Angeles**	7485
11	Long Beach	6710
12	**Antwerp**	6325
13	Qingdao	6307
14	Port Kelang	5544
15	Ningbo	5208
16	Tianjin	4801
17	**New York/New Jersey**	4793
18	Tanjung Pelepas	4169
19	Laem Chabang	3766
20	Tokyo	3594

Source: Port of Hamburg (2005)

shown in Table 6, indicate that the later the year, the greater the AADT, implying that the more recently constructed links carry greater traffic volumes. This finding holds for state routes and US highways, which are both largely products of ad hoc planning, but not for interstate highways, which are more centrally planned, and for which year of construction is insignificant.

4 Model

The case of London rail networks lends credibility to the existence of an inherent first-mover advantage in the development of surface transport networks, and suggests that the advantage derives from spatial location on the network and could be reinforced temporally with increased network connectivity. In order to examine the question of first-mover advantage more rigorously, we proposed an *ex ante* model of network diffusion by which first-mover advantage can be defined and assessed in a controlled spatial environment. Since the purpose of the model is not to be as realistic as possible but to capture the essence of locational and temporal first-mover advantages in a spatial network, we sacrifice some important considerations such as land use development and congestion in order to focus on our research question.

An important simplification of our model is to treat land uses as exogenous and fixed through time. While fully recognizing the impact of land use development on transportation, we are alert to the fact that land use modeling is an intricate process and deserves a separate treatment in its own right. However, it should be noted that, while fixing land uses, this model predicts formation of new places based upon accessibility of potential locations to land use activities. As a network develops, the distribution pattern of accessibility varies accordingly, thereby affecting the formation of new places, and driving a new round of network deployment. In this way, the model partially captures the impact of a transport network on urban growth in a mutual process of network diffusion and place formation. This will be further discussed in the description of the place-formation and link-formation submodels.

Another important simplification deals with link resizing. As a transport network expands, its links (such as roadways and transit lines) and nodes (such as seaports and rail stations) may be resized (generally with increased capacity) to accommodate varying travel demand. In reality, resizing decisions on individual links are made in a complicated investment process that may involve different economic or political initiatives and be limited by the availability of information. To simplify, this study assumes existing links are automatically resized as a network evolves to ensure free-flow travel throughout the network. This assumption is not unreasonable given that our analysis is limited to the early deployment phase of a transport technology,[4] when the issues of congestion and funding deficiency are less significant than during the mature stage. With this assumption, the model eliminates congestion, which otherwise might counteract locational advantages of some heavily used links.

The third simplification is to assume that the advantage of a link or node is proxied by traffic flow traversing that link or node. A link or node with a larger volume of through traffic represents a more critical network element in terms of serving travel needs, improving network connectivity, and increasing surrounding land values. Moreover, given the resizing assumption posited above, a link that carries more traffic is in an advantageous position and will attract a higher level of infrastructure investment. It may be argued that air pollution, visual blight, runoff, and other concerns are serious nuisances associated with traffic flow, but again, in limiting the analysis to the early deployment stage of a network when congestion is not as significant as in the mature stage, this study regards spatially differentiated travel demand as a vital symptom of the locational advantage a facility gains in the network.

4.1 Model Framework

Xie and Levinson (2009c) developed a network growth model called Simulator Of Network Incremental Connection (SONIC), which based network investment decisions on benefit-cost evaluations of potential infrastructure projects. With the assumptions outlined above, this study extends the SONIC model to represent the co-deployment of a surface transport network and places as a bilevel iterative model, which we call SONIC/PF. The outer loop implements a place formation model predicting where a location becomes an established place. The inner loop, on the other hand, includes a simplified travel demand model that predicts traffic flows across an established network and a link formation model that deploys transport links subject to specified economic feasibility criteria to connect established places. The coupled development of places and transport networks distinguishes SONIC/PF from the original SONIC model, which assumed a set of established places at the beginning of network growth.

---

[4] The deployment phase of a transport network is defined as the period when infrastructure is deployed to connect isolated locations as the network expands spatially. It corresponds to the birth and growth stages of the life cycle described in the S-curve theories (Garrison and Levinson 2006; Nakicenovic 1998).

Table 6: Traffic on highways in Minnesota.

	State Highways			US Highways			Interstate Highways		
	Coefficients	*t* Stat	*P*-value	Coefficients	*t* Stat	*P*-value	Coefficients	*t* Stat	*P*-value
Constant	−1 097 804	−3.46	0.0009	−1 267 179	−5.16	0.0000	1 688 215	0.83	0.41
Year	581	3.55	0.0007	674	5.32	0.0000	−810	−0.78	0.44
Adjusted *r*-square	0.14			0.46			0		
N	74			33			29		

The place/link formation model implements a sequential process of place/link addition in an iterative process by which one and only one place is added in an outer-loop round and one and only one link is deployed in an inner round. The process is terminated once candidates are exhausted[5] and the network remains unchanged. The model is illustrated by a flowchart shown in Figure 2, and its component models explained in turn as follows.

Place Formation Model

A place formation model predicts the emergence of new places with a pre-specified distribution of land use activities over an idealized space consisting of a nest of cells. It is assumed that only two types of land use activities exist in the space: labor (housing for workers) and employment (jobs), and both are located at the centroids of the cells.

It is reasonable to posit that a place first forms where desired activities are most accessible, thus we define the locational attractiveness of a centroid in terms of its accessiblity to spatially distributed land use activities. Accessibility is defined as the ease of reaching desired land use activities impeded by the cost of transportation. In this analysis, measures of accessibility adopt a gravity-type form (de Dios Ortuzar and Willumsen 2001). Levinson *et al.* (2007) proposed a composite measure of accessibility that takes into account the accessibility of different types of land use activities. This study considers two major types of accessibility: the ability of a worker to reach jobs across the region, and the ability of an employer to attract a workforce. Assuming one unit of accessibility to workers compensates μ units of accessibility to jobs, the composite measure of accessibility takes the following form:

$$A_i = \mu A_{i,W} + A_{i,J} \tag{1}$$

Where:

$$A_{i,J} = \lg(w_i \sum_j u_j^{-\theta t_{ij}})$$

$$A_{i,W} = \lg(u_i \sum_j w_j^{-\theta t_{ij}})$$

i, j = indices of cell
w_i, u_i = number of workers and number of jobs in cell i
θ = friction factor in the gravity model
t_{ij} = generalized travel time from cell i to cell j

We posit that the potential of a centroid becoming established as a place depends on its accessibility relative to other candidates. The model identifies "local-peak" cells as potential places. A cell is labeled as a local peak when its composite accessibility is greater than all its neighbor cells on the grid. A cell is prohibited from becoming a local peak if it is located on the outer boundary of the region[6] or if any of its neighbor cells has already been established as a place. Local-peak cells make up the choice set, from which one and only one candidate will be selected and established as a new place during each model cycle. The possibility of a local-peak cell becoming established is determined in a logit model depending on its relative accessibility as follows:

$$p_{c_0} = \frac{e^{\eta A_{c_0}}}{\sum_c e^{\eta A_c}} \tag{2}$$

Where:

c = index of local-peak cells
A_c = composite accessibility of candidate c
η = scaling factor indicating how likely a cell with greater composite accessibility gets established)

---

[5] Candidates for a potential place are exhausted when no eligible local peak cells are available; candidates for a potential transport link are exhausted when no potential route could be deployed with a benefit-cost ratio above one. More details will be discussed later.

[6] A cell on the outer boundary would be more likely to become a local peak as it has fewer neighbors. To avoid this source of bias, the place formation model eliminates all the cells on the outer boundary from the candidate set.

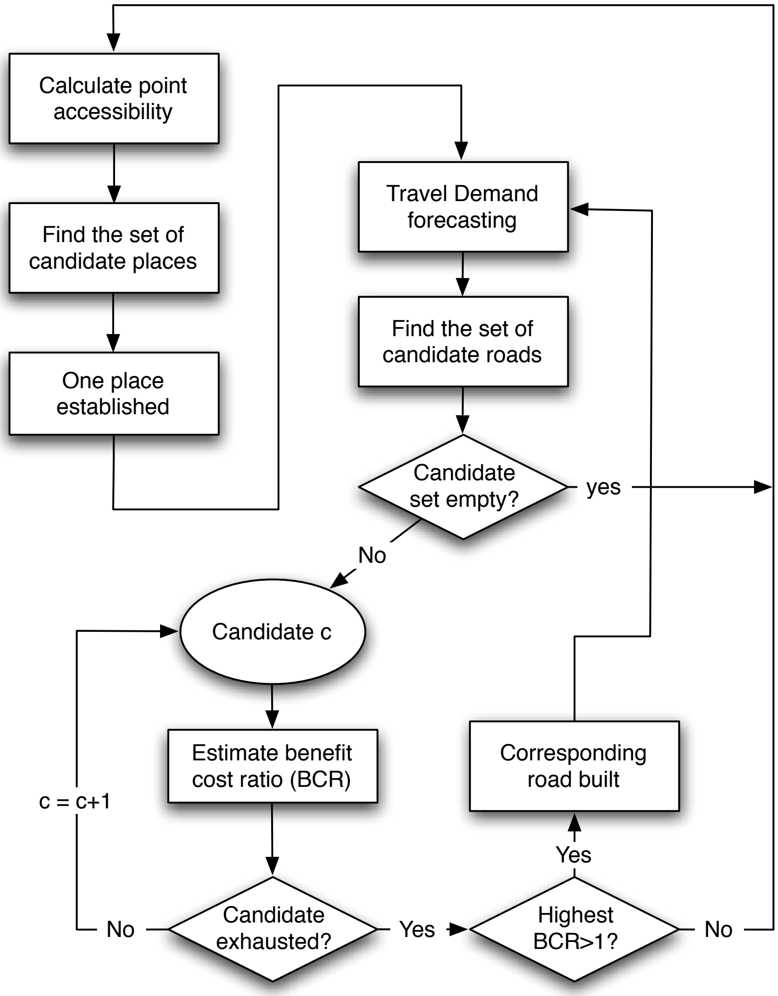

Figure 2: SONIC/PF Model framework

Simplified Travel Demand Model

A simplified travel demand model is proposed to predict traffic flows across an established network; flow volumes are central not only in modeling link formation as discussed later, but also in assessing the locational advantage of a link or node relative to its counterparts in the network. The model includes trip generation, trip distribution, and traffic assignment, while omitting mode choice by assuming a single mode of travel. Trip generation models are made very simple: a worker generates and a job attracts one round trip per day; a doubly-constrained trip distribution model is adopted to predict cell-to-cell trips with the decay factor set as the same as the friction factor of the accessibility models presented above; since no congestion is involved in our model, all-or-nothing traffic assignment is adopted to assign cell-to-cell trips to the lowest-travel-time paths between origins and destinations.

Link Formation Model

A link formation model predicts how transport links are incrementally deployed over space to connect a given set of established places. The model selects one route at a time to build based on benefit-cost evaluations explained as follows:

The benefit of building a potential route is evaluated by the increase in overall accessibility due to the introduction of the proposed route; the cost of deploying a route is estimated by assuming that infrastructure is constructed at a given speed for the same cost rate. Maintenance costs are neglected for simplicity.

Theoretically, infrastructure can be deployed via various routes to connect two places. The path that minimizes travel time and the path that minimizes the map distance (regardless of speeds) between the two places represent two logical options. The former usually maximizes the use of existing links and thus requires less construction while the latter, ignoring the established infrastructure, may require more construction. To limit the number of candidates and reduce the running time, this model only considers these two options for each pair of established places and selects the route that is most cost-effective to build out of all the candidates.

4.2 Simulation Experiments

The model starts with a planar, otherwise undifferentiated space (except as noted below) with neither established places nor transport infrastructure. Land use locations (centroids of land use cells) are connected by primitive trails at a speed of S_l km/h. In a hypothetical scenario as shown in Figure 3 and Figure 4, centroids of land use cells are distributed on a delta grid with the same distance of D kilometers between any pair of neighbor centroids. Each centroid is the center of a hexagonal land use cell, which holds specified numbers of jobs and workers, both assumed to be fixed over time. Christaller (1933); King (1985) demonstrated in central place theory that activities are distributed at nodes of different levels in the hexagonal network, which represent centers of nested hexagons. In this case, centroids with the distance of D kilometers belong to the lowest level, centroids with the distance of $2D$ belong to the second level, etc. The local-peak assumption of this model essentially requires that a centroid be classified in the second level or higher to qualify as a candidate place. The value of a one-unit increase in accessibility is monetized as $\$v$ and remains fixed over time. A transport link is deployed with a uniform design speed of S_h km/h on top of a trail, for a constant cost of $\$C$ per kilometer.

While land uses are exogenous and fixed over time, two different initial land use distributions are tested to examine the sensitivity of our analysis to the land use inputs. Two experiments are executed accordingly: in Experiment A, the numbers of jobs and workers in each cell are randomly allocated; in Experiment B, the number of jobs in each cell declines exponentially at a rate of β_1 with increasing distance between the cell and the center of the space, while the number of workers increases exponentially at a rate of β_2. In both experiments, the total number of workers is assumed to equal the total number of jobs and the average number of jobs or workers in a cell is fixed at Q. Table 7 lists the default values of coefficients and parameters set in the model.

4.3 Hypotheses

Now that the model is set up to simulate the spatial development of a transport network, it can be employed to test hypotheses regarding the locational and temporal advantages of first-mover places or links in the network. Imagine an extreme case in which initial land uses are highly concentrated: pivotal locations where settlements are concentrated are likely to be established first; then transport facilities are built to connect these places, and become strategic routes that are expected to carry high volumes of through traffic. At this stage, it is posited that earlier-established places and transport facilities would gain FMA simply because they have acquired the best locations.

As the network spreads and connects to smaller places, it brings more traffic to earlier-established places and strategic links. This network effect is expected to reinforce the first-mover advantage during the evolutionary process of network growth.

The advantages of first-movers would be less salient, however, if:

1. the initial land use distribution is less concentrated (if a place forms being at least paramount to earlier-established places, it may divert trips from their original destinations and undermine the advantages of established places and transport facilities serving them), or;

2. the network is over-invested (if multiple routes are built between the same origin and destination, routes may compete with each other for the travel demand, thereby reducing the dominance of earlier-deployed routes).

Based on these speculations, the following hypotheses are proposed and tested in simulation experiments:

H1: Earlier-established places and transport links gain FMA in a network, which will be reinforced as the network grows over time.

H2: FMA is less evident in Experiment A than in Experiment B, as the latter represents a greater concentration of land uses.

H3: FMA is less evident in a more redundant network, as it indicates more intensive competition between parallel routes.

4.4 Measurement

At the end of each inner-loop iteration, the model outputs the formation time of each established place and transport link,[7]

---

[7] Only one place is established in a outer-loop iteration, so its formation time is indicated by iteration number. For links, their formation times are also distinguished by their order in the sequence of construction. For instance, a link formation time labeled as "16.02" indicates this link is the second link that is built in Iteration 16.

Table 7: Specified values of model parameters.

Para.	Value	Unit	Description
θ	0.05	/min	Decay factor in trip distribution and the friction factor in node formation
μ	1	N.A.	Relative value of accessibility to workers compared to accessibility to jobs
η	3	N.A.	Scaling factor in node formation model
β_1	0.15	/km	Decay rate of jobs from region center
β_2	0.05	/km	Increase rate of workers from region center
C	1000000	\$/km	Construction cost of paved roads
D	5	km	Distance between adjacent land use centroids
Q	500	N.A.	Average number of jobs or workers in a land use cell
S_l	10	km/hr	Specified uniform speed of primitive trails
S_h	30	km/hr	Specified uniform speed of transport links
v	0.05	\$/unit	Monetary value of a unit of accessibility to jobs

as well as traffic volumes entering each place and each link. If FMA does exist, earlier-established places and links should attract more traffic. The relationship between the ranks of places or links in terms of their formation times and those in term of their traffic flows is examined by the Spearman rank order correlation test (Higgins 2003), a non-parametric measure of correlation assessing how well an arbitrary monotonic function describes the relationship between two variables without making any assumptions about the frequency distribution of the variables. A negative Spearman correlation coefficient would indicate the presence of FMA, suggesting that the earlier a place or link is established, the larger volume of traffic it attracts; a positive correlation coefficient would indicate a first-mover disadvantage. The absolute value of the correlation coefficient would indicate the significance of the first-mover advantage or disadvantage.

In order to test the relationship between FMA and network redundancy, this study proposes two topological measures. The first is the γ index, a connectivity measure that quantifies the interconnection of nodes in a network (Harggett and Chorley 1969) by comparing the actual number of links with the maximum number of possible links in the network:

$$\gamma = \frac{e}{6(v-2)} \tag{3}$$

Where
e= number of directional edges, and
v= number of vertices (nodes).

The second measure, "circuitness," is adopted from Xie and Levinson (2007a), who developed an algorithm to identify the predefined structural elements of ring, web, circuit, and branch in a network and evaluate their relative significance according to link lengths. If a link is located on one and only one circuit, it belongs to a ring; if it is located on more than one circuit, it belongs to a web. If a link belongs to a web or ring, it is defined as a *circuit link*; otherwise, it is defined as a *branch link*. Therefore,

$$\phi_{\text{circuit}} = \phi_{\text{ring}} + \phi_{\text{web}} \tag{4}$$

Where

$$\phi_{ring} = \frac{\sum_i (l_i \delta_i^{ring})}{\sum_i l_i}$$

$$\phi_{web} = \frac{\sum_i (l_i \delta_i^{web})}{\sum_i l_i}$$

l_i= length of an individual edge i
δ_i^{ring}= 1 if link i belongs to a ring; 0 otherwise
δ_i^{web}= 1 if link i belongs to a web; 0 otherwise

4.5 Results

Experiment A stopped at the twenty-ninth iteration and Experiment B terminated at the twenty-seventh iteration. Figures 3 and 4 display the snapshots of the evolving network in the two experiments, respectively. Gray dots represent the centroids of land use cells, some of which change to magenta when established as places. The relative size of a dot indicates the agglomeration scale of land use activities (workers plus jobs) at a specific location. Gray edges represent primitive trails, some of which change to blue when they are built as transport links.

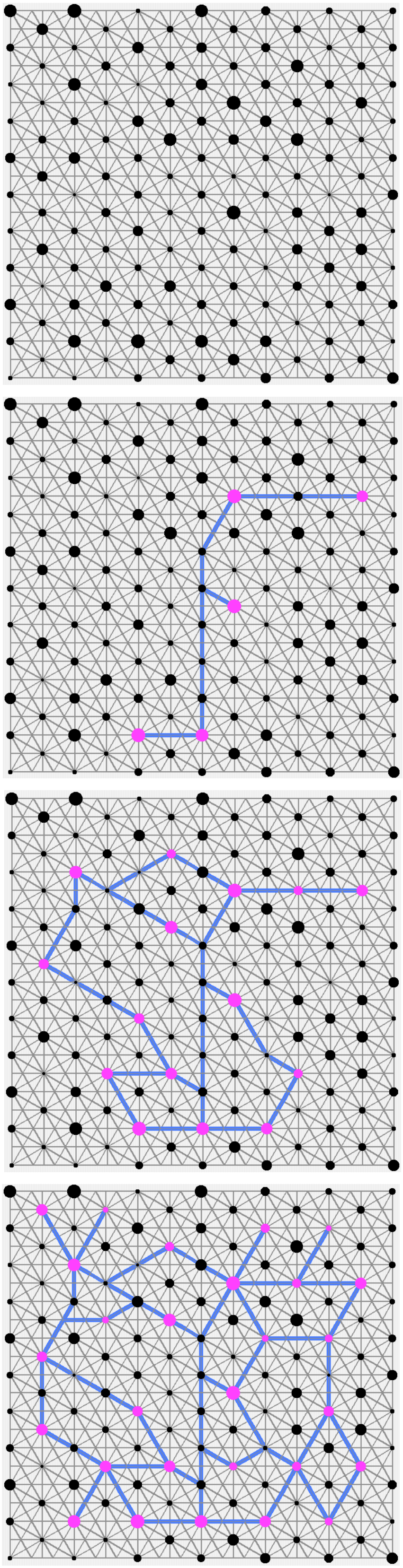

Figure 3: Snapshots at Iteration 0, 5, 15 and 29 in Experiment A.

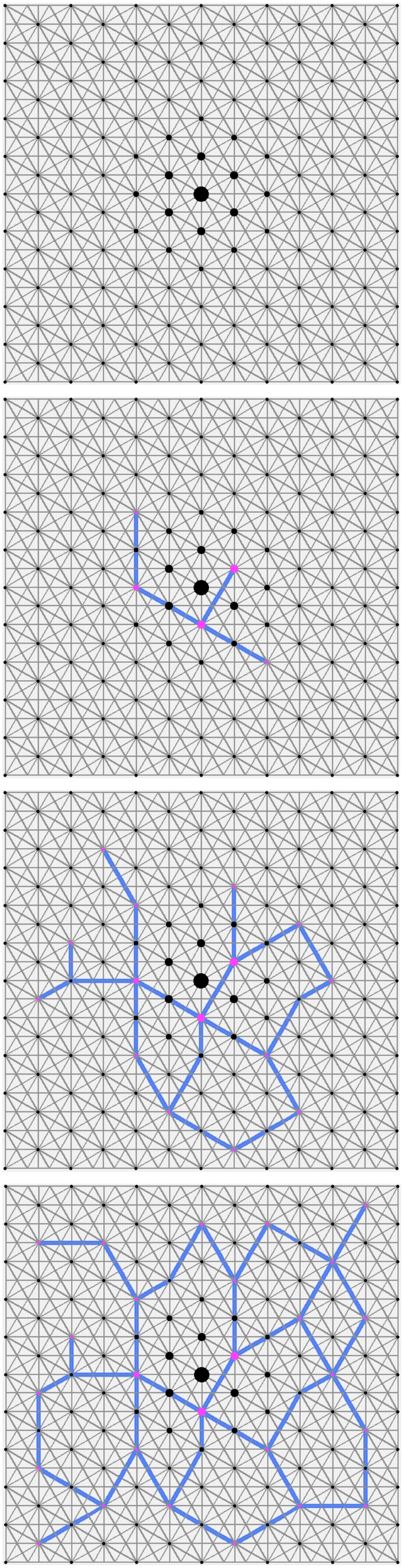

Figure 4: Snapshots at Iteration 0, 5, 15 and 27 in Experiment B.

Spearman correlation tests were carried out for both places and transport links at the end of every other iteration, and the proposed topological measures were computed as well. Tables 8 and 9 present the results from Experiments A and B, respectively. The fluctuations of correlation coefficients and topological measures over iterations are displayed in Figures 5 and 6, respectively. Only the correlation coefficients with a 90 percent or higher confidence level (i.e., $p < 0.10$) are presented.

In most cases, the correlation between formation time and traffic volume is negative for both places and links, suggesting that the earlier a place or a link is established, the more traffic it attracts. This provides evidence for the existence of FMA in the deployment of a network. A general trend of increase in the absolute value of the correlation coefficient for both places and links over time is also observed, suggesting that FMA is reinforced as the network expands.

Starting with a more concentrated bell-shaped distribution of land uses, Experiment B results in stronger negative Spearman correlations than Experiment A, suggesting a more concentrated distribution of land uses leads to more significant first-mover advantages in the formation of a transport network serving these land uses.

Both topological measures (the γ index and the measure of circuitness) indicate the generally increasing redundancy of the simulated network. The fluctuation of the circuitness measure is more volatile as compared to that of the γ index. The rises on the circuitness curve indicate the additions of circuit links that create alternative routes, while the falls reflect the addition of branch links. Interestingly, as can be seen in Iterations 11–17 in Experiment A and Iterations 7–9 and 17–21 in Experiment B, the increase in circuitness is always accompanied by the weakening of the Spearman correlation. This observation suggests an inherent correlation between FMA and network redundancy, as posited in the third hypothesis—although rigorous statistical tests are still needed to substantiate this relationship.

4.6 Sensitivity Analysis

The values of model parameters listed in Table 2 are arbitrarily specified. To test the sensitivity of our analysis to these parameters, simulation was re-executed in a series of model runs in which the values of each parameter were altered. The results are summarized in Table 10.

A smaller decay factor θ in the gravity model indicates a smaller impedance across a network and a higher level of accessibility. As can be seen, a smaller decay factor (0.02) in Run 1 for Experiment A resulted in much smaller Spearman correlation coefficients for both places (-0.482) and links (-0.209), indicating weaker first-mover advantages. This agrees with the speculation that advantages of first movers deriving from locational advantage in a network will be undermined as falling travel impedence reduces locational differentiation.

A smaller scaling factor η allows more randomness in the formation of places, thereby counteracting the first-mover advantages. Similarly, a smaller value of μ or β_2 (in the bell-shaped distribution of land uses), specifying a lower concentration of initial land uses, is expected to lead to smaller first-mover advantages as well. To test this, Experiment B was rerun in Run 2 with a different value of β_2 (0.10), producing a weaker (and statistically significant) correlation for both places (-0.760) and links (-0.654).

A lower value for accessibility (v) or a higher construction cost rate (C) leads to less construction in general, because the link formation process considers both benefit and cost. As the result of less network redundancy, more evident FMA is expected to be observed. Experiment A was re-run in Run 3 with a different value of C (500 000), and a stronger FMA for links (-0.611) was observed.

The distance between adjacent centroids D indicates the magnitude of the space and network, while Q indicates the scale of land use agglomerations. Changing either variable with the other remaining equal would change land uses and travel needs. Experiment A was re-run in Run 4 with a different value of Q (1000). The resulting Spearman correlation coefficients were -0.609 for places and -0.581 for links, indicating a slightly weaker FMA for places and a slightly stronger FMA for links.

The higher the design speed for transport links, the faster one can travel across established transport links versus primitive trails, and a stronger FMA is expected in more strongly differentiated networks. Re-running Experiment A with a higher speed (60), as expected, resulted in a much stronger correlation for both places (-0.710) and links (-0.733).

5 Discussion

First-mover advantages depend on several network characteristics.

First, are we considering nodes or links? This paper examines both. A node can connect to many links, while a link can connect to only two nodes, so we expect that first-mover effects for nodes and links will be different. The capacity of nodes and links may be considered in different ways. Nodes may have limits on number of vehicles (flow) or on number of incoming or outgoing links (capacity). Similarly, links may

Table 8: Topological measures and Spearman correlation coefficients computed in Experiment A.

			Node		Link	
Iteration	Circuitness	Gamma	Coeff.	P-value	Coeff.	P-value
1	0.000	0.349	0.000	0.992	N.A.	
3	0.000	0.346	0.500	0.478	0.298	0.066
5	0.000	0.344	0.100	0.834	-0.080	0.992
7	0.000	0.351	-0.214	0.596	-0.242	0.317
9	0.391	0.348	0.067	0.849	-0.549	0.002
11	0.329	0.350	-0.300	0.342	-0.539	0.000
13	0.593	0.354	-0.379	0.187	-0.451	0.000
15	0.911	0.357	-0.318	0.234	-0.465	0.000
17	1.000	0.359	-0.127	0.610	-0.430	0.000
19	0.967	0.358	-0.393	0.095	-0.455	0.000
21	0.936	0.356	-0.495	0.026	-0.448	0.000
23	0.880	0.358	-0.557	0.009	-0.499	0.000
25	0.859	0.362	-0.642	0.002	-0.577	0.000
27	0.901	0.361	-0.562	0.004	-0.611	0.001
29	0.879	0.361	-0.611	0.001	-0.567	0.000

Table 9: Topological measures and Spearman correlation coefficients computed in Experiment B.

			Node		Link	
Iteration	Circuitness	Gamma	Coeff.	P-value	Coeff.	P-value
1	0.000	0.364	0.000	0.992	N.A.	
3	0.000	0.356	0.500	0.478	0.563	0.038
5	0.000	0.346	0.500	0.312	0.460	0.052
7	0.000	0.351	-0.500	0.219	-0.390	0.063
9	0.539	0.353	-0.250	0.478	-0.382	0.021
11	0.735	0.350	-0.473	0.134	-0.469	0.001
13	0.647	0.353	-0.813	0.005	-0.586	0.000
15	0.720	0.351	-0.804	0.003	-0.680	0.000
17	0.654	0.352	-0.887	0.000	-0.786	0.000
19	0.693	0.356	-0.809	0.001	-0.777	0.000
21	0.913	0.354	-0.808	0.000	-0.786	0.000
23	0.855	0.357	-0.896	0.000	-0.810	0.000
25	0.896	0.357	-0.822	0.000	-0.791	0.000
27	0.872	0.357	-0.816	0.000	-0.757	0.000

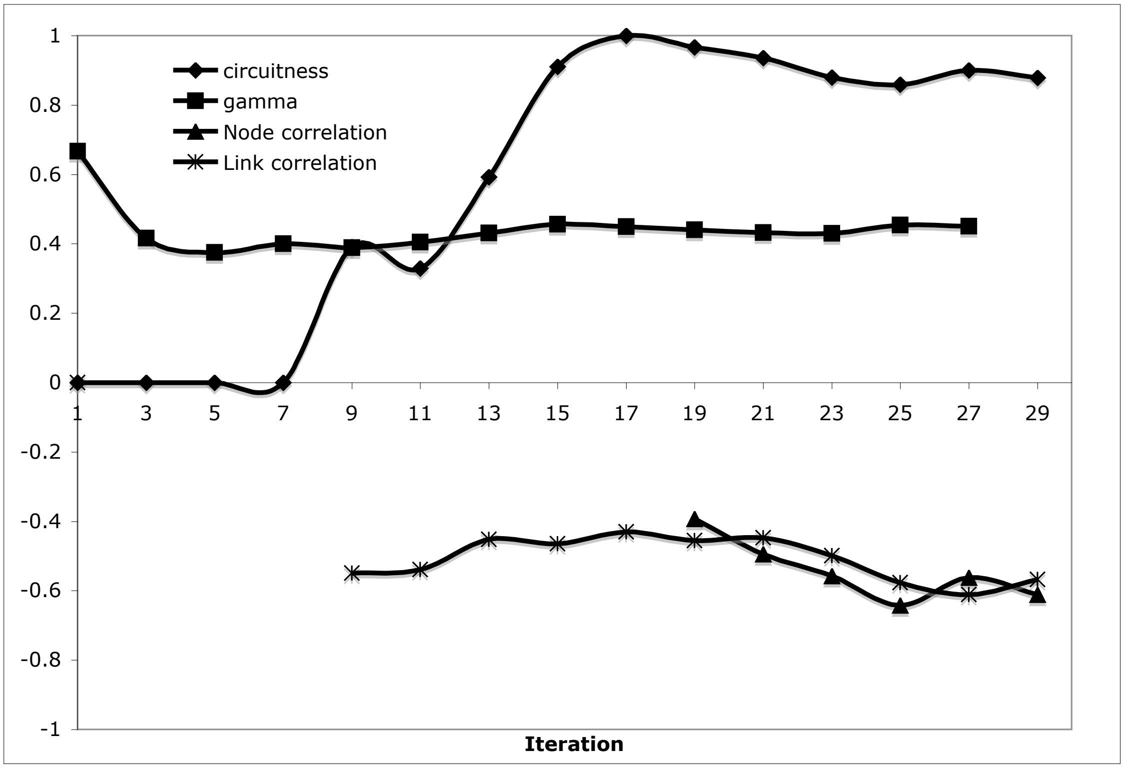

Figure 5: The temporal change of topological attributes and Spearman rank order correlation in Experiment A.

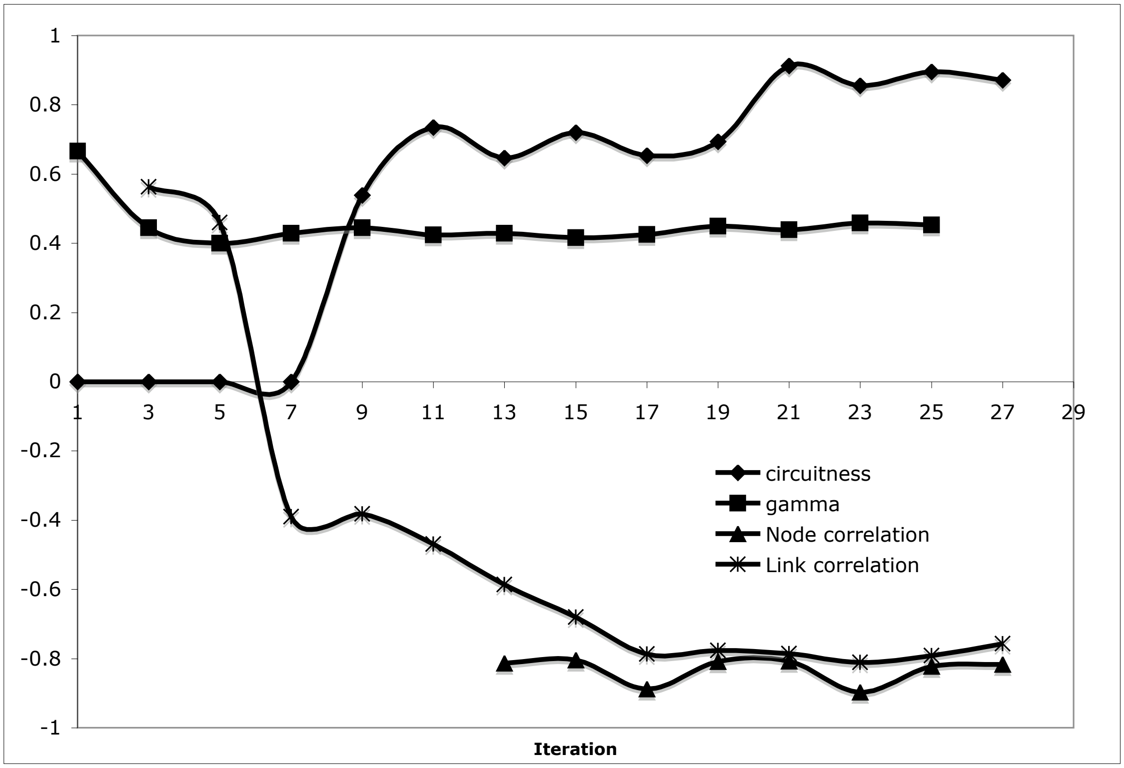

Figure 6: The temporal change of topological attributes and Spearman rank order correlation in Experiment B.

Table 10: Spearman correlation coefficients in sensitivity analysis.

		Value		Node		Link	
Run	Para.	Previous	Current	Coeff.	p-value	Coeff.	p-value
1	θ	0.05	0.02	-0.482	0.018	-0.209	0.016
2	β_1	0.15	0.1	-0.760	0.000	-0.654	0.000
3	C	1 000 000	5 000 000	-0.389	0.047	-0.611	0.000
4	Q	500	1000	-0.609	0.002	-0.581	0.000
5	S_h	30	60	0.710	0.000	-0.733	0.000

also have a flow-defined capacity limit, or it may be limited in the number of lanes. Since nodes can connect to more links than links can connect to nodes, we expect nodes to be more eligible for FMA than links.

Second, is there a preference for attaching to existing network elements in a particular way? Nodes may benefit from preferential attachment (Newman 2001), while links benefit from preferential reinforcement (Yerra and Levinson 2005), where existing links with large capacities attract more investment. There are both supply-side and demand-side reasons for these preferences. Supply-related causes include economies of scale, economies of density, and lack of capacity constraint. Demand-related causes include network effects. Preferential attachment favors FMA.

Third, are we considering capacity-constrained or capacity-unconstrained networks? (This paper considers unconstrained networks) All networks are ultimately constrained, but if the network in question is (for practical purposes) unconstrained, we get different answers than when dealing with a congested network. Unconstrained networks are more likely to exhibit FMA.

Fourth, are there network externalities? When network externalities are present, there is an advantage to hubbing. However, as capacity constraints are approached, congestion externalities present a disadvantage to hubbing. The net effect depends on the technological characteristics of the mode as well as demand conditions. If hubbing benefits exceed congestion costs, then first-mover advantages are possible. The London Underground and the hypothetical uncongested road network both illustrate FMA in transportation networks. The international system of airports is not subject to FMA—the first airports do not carry more traffic than later airports. The international system of seaports also do not possess FMA. However, the location of hub cities within an airline system is persistent. Airlines maintain hubs in the cities where they were first established.

Fifth, are coordination advantages spatial, temporal, or both? Fixed infrastructure is spatially coordinated, while transportation services (carriers such as airlines, shippers, buses, etc.) are coordinated both spatially and temporally, and so has greater potenital for coordination economies. For example, the greatest spatial improvement (distance reduction) for a road network over a standard grid is circuity, which is on the order of 20 percent distance savings for a true air-line connection rather than a more typical network connection (Levinson and El-Geneidy 2009). Speeds may change as well, though.

For a carrier network with scheduled services, hubbing can reduce schedule delays significantly by concentrating sufficient demand. Because the network economies are greater at hubs, the first hub (particularly if it is served by multiple carriers) has a greater advantage over later hubs.

6 Conclusions

This paper investigates the existence and extent of first-mover advantages in the deployment of spatial surface transport networks. Examining the case of London railroads suggests inherent first-mover advantage in a surface transport network, and indicates that the advantage derives from spatial location and could be reinforced temporally with increased network connectivity. A network diffusion model is then developed to replicate the growth of transport networks over space and time, to test if earlier-established places and transport facilities gain locational advantages, and to determine if the advantages remain the same or change during the evolutionary process of network growth. Using traffic flow as a proxy for locational advantage in the early deployment phase of a network, Spearman rank order correlation tests reveal that the earlier a place or a link is established, the larger the volume of traffic it attracts; the finding that the correlation becomes stronger as the network grows suggests that first-mover advantages not only exist in transport networks, but are reinforced as the network expands. Simulation results also reveal that the extent of first-mover advantages in a transport network correlates with initial land use distribution and network redundancy.

In contrast to the game-theoretic methods widely adopted in previous FMA studies, this research contributes to the literature by proposing a modeling approach in which first-mover advantage is defined and analyzed in a controlled environment. Although this study sacrifices some important considerations regarding land development, congestion, ownership, and investment decision-making, it keeps the model simple to examine the particular question of first-mover advantages. Elsewhere, the authors have treated other matters in a series of parallel studies on network growth. Under the umbrella of network growth, the authors have conducted separate studies to examine the co-development of transportation and land use using the empirical data from the Minneapolis-Saint Paul streetcar system (Xie and Levinson 2009b) and to model the coupled development in an autonomous process (Levinson *et al.* 2007). The authors have also constructed a theoretic model to analyze the relationship between transport infrastructure and its governing agencies (Xie and Levinson 2009a).

As evidence has revealed the existence of first-mover advantages in the deployment of surface transport networks, this research has important implications for strategic transport planning, investment, and network design. The builders of transport networks need to be exceedingly careful that the networks are appropriately sized and sited, since these decisions will shape the use of those networks profoundly as the system adapts and locks in. In addition, there are many research questions yet to answer: How can economic and political initiatives factor into the deployment of a transport system? How should transportation funds be allocated between existing infrastructure and new construction to facilitate the growth of a region? How should a transportation facility be appropriately sized and sited if the goal is not necessarily to optimize it for current conditions but to improve the system as a whole, considering future construction? It may not be possible to answer some of these questions without developing a more sophisticated network model. This study, however, serves as a starting point in that it recognizes the existence of first-mover advantage and proposes a network diffusion model to investigate the factors contributing to it; the model presented here has the potential to serve as a planning tool that takes into account the effects of first-mover advantages.

Acknowledgments

The authors sincerely thank the two anonymous reviewers who provided constructive comments on this paper. This research has been supported by the UK Economic and Social Research Council. This material is based upon work supported by the United States National Science Foundation under Grant No. 0236396. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

About the authors

David Levinson (dlevinson@umn.edu) is the R.P. Braun Chair in Transportation Engineering and Associate Professor of Civil Engineering at the University of Minnesota, Minneapolis. He is the editor of the *Journal of Transport and Land Use* and director of the Networks, Economics, and Urban Systems (NEXUS) research group (http://nexus.umn.edu).

Feng Xie (fxie@mwcog.org) is a Transportation Engineer with the Metropolitan Washington Council of Governments. He received his PhD in Civil Engineering from the University of Minnesota in 2008.

References

Airports Council International. 2006. Traffic movements 2005. Technical report, Airports Council International.

Alberti, M. and P. Waddell. 2000. An integrated urban development and ecological simulation model. *Integrated Assessment*, 1:215–227.

Anas, A. 2004. Vanishing cities: What does the new economic geography imply about the efficiency of urbanization? *Journal of Economic Geography*, 4:181–199. doi: 10.1093/jeg/4.2.181.

Anas, A. and R. J. Arnott. 1993. Development and testing of the Chicago Prototype Housing Market Model. *Journal of Housing Research*, 4(1):73–130.

Anas, A. and Y. Liu. 2007. A regional economy, land use, and transportation model (RELU-TRANS): Formulation, algorithm design, and testing. *Journal of Regional Science*, 47(3):415–455. doi: 10.1111/j.1467-9787.2007.00515.x.

Barrat, A., M. Barthélemy, and A. Vespignani. 2004. Modeling the evolution of weighted networks. *Physical Review E*, 70(6):66–149. doi: 10.1103/PhysRevE.70.066149.

Bates, J., M. Brewer, P. Hanson, D. McDonald, and D. Simmonds. 1991. Building a strategic model for Edinburgh.

In *Proceedings of Seminar D, PTRC 19th Summer Annual Meeting. PTRC, London*.

Borley, H. 1982. *Chronology of London Railways*. Oakham, UK: Railway and Canal Historical Society.

Boyce, D. 2007. An account of a road network design method: Experssway spacing, system configuration and economic evaluation. In *Infrastructure Problems under Population Decline*, pp. 1–30. Berlin: Berliner Wissenschafts-Verlag.

BTNews Online. 2004. Siegel: Philadelphia could be US Airways' last stand. *Business Travel News Online*. URL http://www.btnmag.com/businesstravelnews/headlines/article_display.jsp?vnu_content_id=1000471812.

Center for Economic and Social Inclusion. 2006. National statistics: First release: Labour market statistics. Technical report, Center for Economic and Social Inclusion.

Christaller, W. 1933. *Die zentralen Orte in Suddeutschland*. Jena: Gustav Fischer.

Corbett, M., F. Xie, and D. Levinson. 2008. Evolution of the second-story city: The Minneapolis Skyway system. *Environment and Planning, Part B*, 36(4):711–724. doi: 10.1068/b34066. URL http://nexus.umn.edu/Papers/Skyways.pdf.

Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein. 1990. *Introduction to Algorithms*. The MIT Press.

de Dios Ortuzar, J. and L. G. Willumsen. 2001. *Modeling Transport*. John Wiley and Sons.

Diers, J. and A. Isaacs. 2006. *Twin Cities by Trolley: The Streetcar Era in Minneapolis and St. Paul*. University of Minnesota Press.

Fujita, M. and P. Krugman. 2003. The new economic geography: Past, present and the future. *Papers in Regional Science*, 83:139–164. doi: 10.1007/s10110-003-0180-0.

Garrison, W. and D. Levinson. 2006. *The Transportation Experience: Policy, Planning, and Deployment*. Oxford University Press.

Gibbons, R. 1992. *Game Theory for Applied Economists*. Princeton University Press.

Hansen, W. 1959. How accessibility shapes land use. *Journal of American Institute of Planners*, 25:73–76.

Harggett, P. and J. C. Chorley. 1969. *Network Analysis in Geography*. Butler and Tanner Ltd.

Helbing, D., J. Keltsch, and P. Moln·r. 1997. Modeling the evolution of human trail systems. *Nature*, 388:47.

Higgins, J. J. 2003. *Introduction to Modern Nonparametric Statistics*. Duxbury Press.

Iacono, M., D. Levinson, and A. El-Geneidy. 2008. Models of transportation and land use change: A guide to the territory. *Journal of Planning Literature*, 22:323–340. doi: 10.1177/0885412207314010. URL http://nexus.umn.edu/Papers/MTLUC.pdf.

Kerin, R., P. Varadarajan, and R. Peterson. 1996. First-mover advantage: A synthesis, conceptual framework, and research propositions. *Journal of Marketing*, 56(4):33–52.

King, L. 1985. *Central Place Theory*. London: SAGE Publications Ltd.

Kondratieff, N. 1987. *The Long Wave Cycle*. New York: Richardson and Snyder.

Krugman, P. 1992. *Geography and Trade*. The MIT Press.

Lachene, R. 1965. Networks and the location of economic activities. *Regional Science Association, Papers*, 14:183–196.

Lam, L. and R. Pochy. 1993. Active-walker models: Growth and form in nonequilibrium systems. *Computation Simulation*, 7:534.

LeBlanc, L. J. 1975. An algorithm for the discrete network design problem. *Transportation Science*, 9(3):183–199.

Levinson, D. 2008a. Density and dispersion: The codevelopment of land use and rail in London. *Journal of Economic Geography*, 8(1):55–57. doi: 10.1093/jeg/lbm038. URL http://nexus.umn.edu/Papers/Codeploy.pdf.

Levinson, D. 2008b. The orderliness hypothesis: Does population density explain the sequence of rail station opening in London? *Journal of Transport History*, 29(1):98–114.

Levinson, D. and A. El-Geneidy. 2009. The minimum circuity frontier and the journey to work. *Regional Science and Urban Economics*, 39(6):732–738. doi: 10.1016/j.regsciurbeco.2009.07.003. URL http://nexus.umn.edu/Papers/Orderliness.pdf.

Levinson, D., F. Xie, and S. Zhu. 2007. The co-evolution of land use and road networks. In *Proceedings of the 17th International Symposium on Transportation and Traffic Theory (ISTTT)*. URL http://nexus.umn.edu/Papers/SIGNAL2007.pdf.

Levinson, D. and B. Yerra. 2006. Self organization of surface transportation networks. *Transportation Science*, 40(2):179–188. doi: 10.1287/trsc.1050.0132. URL http://nexus.umn.edu/Papers/SelfOrganization.pdf.

Levinson, M. 2006. *The Box*. Princeton University Press.

Lieberman, M. and D. Montgomery. 1988. First-mover advantages. *Strategic Management Journal*, 9:41–58. doi: 10.1002/smj.4250090706.

Lieberman, M. and D. Montgomery. 1998. First-mover (dis) advantages: Retrospective and link with the resource-based view. *Strategic Management Journal*, 19(12):1111–1125. doi: 10.1002/(SICI)1097-0266(1998120)19:12<1111::AID-SMJ21>3.0.CO;2-W.

Mackett, R. 1983. The Leeds Integrated Transport Model (LILT). Supplementary Report 805, Transport and Road Research Laboratory, Crowthorne, UK.

Makadok, R. 1998. Can first-mover and early-mover advantages be sustained in an industry with low barriers to entry/imitation? *Strategic Management Journal*, 19(7):683–696. doi: 10.1002/(SICI)1097-0266(199807)19:7<683::AID-SMJ965>3.0.CO;2-T.

Marlette, J. 1959. *Electric Railroads of Indiana*. Indianapolis: Council for Local History.

Mittal, S. and S. Swami. 2004. What factors influence pioneering advantage of companies? *The Journal for Decision Makers*, 29(3):15–33.

Mueller, D. 1997. First-mover advantages and path dependence. *International Journal of Industrial Organization*, 15(6):827–850. doi: 10.1016/S0167-7187(97)00013-1.

Nakicenovic, N. 1998. Dyanmics and replacement of U.S. transport infrastructure. In *Cities and Their Vital Systems: Infratructure, Past, Present and Future*. Washington, D.C.: National Academy Press.

Newman, M. 2001. Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2):25102. doi: 10.1103/PhysRevE.64.025102.

Newman, M. 2003. The structure and function of complex networks. *SIAM Review*, 45:167–256. URL http://www.jstor.org/stable/25054401.

Port of Hamburg. 2005. Top 20 ports: Container throughput in TEU. Technical report, Port of Hamburg.

Pred, A. 1966. *The Spatial Dynamics of U.S. Urban-Industrial Growth, 1900–1914*. Cambridge, MA: The MIT Press.

Puffert, D. 2002. Path dependence in spatial networks: The standardization of railway track gauge. *Explorations in Economic History*, 39(3):282–314. doi: 10.1006/exeh.2002.0786.

Rahman, Z. and S. Bhattacharyya. 2003. First mover advantages in emerging economies: A discussion. *Management Decision*, 41(2):141–147.

Rimmer, P. 1967. The changing status of New Zealand seaports. *Annals of the Association of American Geographers*, 57:88–100.

Rose, D. 1983. *The London Underground: A Diagrammatic History*. Douglas Rose.

Rosenstein, M. 2000. *The Rise of Maritime Containerization in the Port of Oakland: 1950 to 1970*. Master's thesis, Gallatin School of Individualized Study, New York University.

Streeter, W. 1946. Is the first move an advantage? *Chess Review*, p. 16.

Taaffe, E., R. L. Morrill, and P. R. Gould. 1963. Transportation expansion in underdeveloped countries: A comparative analysis. *Geographical Review*, 53(4):503–529. URL http://www.jstor.org/stable/212383.

Timmermans, H. 2003. The Saga of Integrated Land Use-Transport Modeling: How Many More Dreams Before We Wake Up? In *Proceeding of 10th International Conference on Travel Behaviour Research, Lucerne, August 10–15, 2003*.

von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

von Thünen, J. H. 1910. *Isolated State*. Jena: Gustav Fischer.

Xie, F. 2008. Validation of the model of network degeneration: A case study of the Indiana interurban network. (08-0959) Presented at 87th annual meeting of the Transportation Research Board (TRB), Washington, D.C.

Xie, F. and D. Levinson. 2007a. Measuring the structure of road networks. *Geographical Analysis*, 39(3):336–356. doi: 10.1111/j.1538-4632.2007.00707.x. URL http://nexus.umn.edu/Papers/Topology.pdf.

Xie, F. and D. Levinson. 2007b. The weakest link: A model of the decline of surface transportation networks. *Transportation Research E*, 44(1):100–113. doi: 10.1016/j.tre.2006.09.001. URL http://nexus.umn.edu/Papers/WeakestLink.pdf.

Xie, F. and D. Levinson. 2009a. Governance choice on a serial network. *Public Choice*, 141(1). doi: 10.1007/s11127-009-9448-5. URL http://nexus.umn.edu/Papers/GovernanceChoice.pdf.

Xie, F. and D. Levinson. 2009b. How streetcars shaped suburbanization: A Granger-causality analysis of land use and transit in the Twin Cities. *Journal of Economic Geography*, 10(3). doi: 10.1093/jeg/lbp031. URL http://nexus.umn.edu/Papers/Streetcar.pdf.

Xie, F. and D. Levinson. 2009c. Jurisdictional control and network growth. *Networks and Spatial Economics*, 9(3):459–483. doi: 10.1007/s11067-007-9036-5. URL http://nexus.umn.edu/Papers/SONIC.pdf.

Xie, F. and D. Levinson. 2009d. Modeling the growth of transportation networks: A comprehensive review. *Networks and Spatial Economics*, 9(3):291–307. doi: 10.1007/s11067-007-9037-4. URL http://nexus.umn.edu/Papers/ReviewOfNetworkGrowth.pdf.

Yamins, D., S. Rasmussen, D., and Fogel. 2003. Growing urban roads. *Networks and Spatial Economics*, 3:69–85. doi: 10.1023/A:1022001117715.

Yang, H. and M. G. H. Bell. 1998. Models and algorithms for road network design: A review and some new developments. *Transport Reviews*, 18(3):257–278. doi:

10.1080/01441649808717016.

Yerra, B. and D. Levinson. 2005. The emergence of hierarchy in transportation networks. *Annals of Regional Science*, 39(3):541–553. doi: 10.1007/s00168-005-0230-4. URL http://nexus.umn.edu/Papers/Emergence.pdf.

Zhang, L. and D. Levinson. 2004. A model of the rise and fall of roads. Presented at Engineering Systems Symposium, Massachussetts Institute of Technology. URL http://nexus.umn.edu/Papers/RiseAndFallOfRoads.pdf.