

# Geographic Variation in Internet Connectivity

Luke M. Froeb, Richard W. Oliver, David A. Weiskopf

Vanderbilt University  
Nashville, TN

## Abstract

*This paper presents a measure of Internet connectivity" (number of connected companies per capita) normalized by the economic "size " of areas in the United States. H"e use cross sectional variation in connectivity to test some h5potheses about technology adoption. We find that areas in the South and Midwest have below average connectivity. Connectivity is also related to the living "attractiveness " ofan area*

## 1. Introduction

Dramatic technological advances in telecommunications, particularly the development ofthe Internet, are going to have equally dramatic effects on the evolution of commerce. There is no shortage of forecasts, but rarely are such forecasts informed by data. Part of the problem is that forecasts are necessarily forward looking and the explosive growth of the Internet is a relatively recent phenomenon. This makes data scarce by shortening the relevant historical window from which to generalize.

This paper utilizes cross sectional variation in Internet connectivity to test some hypotheses about technology adoption, and to provide some stylized observations that can be used to inform future forecasts. By focusing on economic areas we are able to address one of the major concerns, that the development of the Internet is going to make cities obsolete by destroying their economic *raison detre*.

We examine what we call the "market widening" and "firm devolution" effects of the Internet in loosening geographic proximity constraints faced by consumers and workers. If the bandwidth revolution lives up to its promise, no longer will consumers and employees be constrained to live in close proximity to retail stores or workplaces. These changes suggest a "de-urbanization" or at least a movement towards more attractive cities, that

will partially reverse the effects of the urban migrations of the last century.

In what follows, we explore these hypotheses in greater detail by first examining the reasons for the growth of cities as a means to shed light on their possible demise. For certain types of goods and services, we see a "market widening" effect--distant firms serving separate geographic markets will be brought into competition with one another. Firms will serve wider geographic markets and become more specialized with the net effect that consumers are offered a wider array of consumption choices.

We then examine what we call the "firm devolution" effects of the Internet. Firms will move certain within-firm transactions out from under the corporate umbrella and into the marketplace. We examine these hypotheses with an eye towards developing testable implications. We then present our measure of connectivity and relate it to the hypotheses examined. We close with conclusions and suggestions for future empirical work.

## 2. Market Widening

The best way to understalid the potential demise of cities is to examine the economic reasons for their formation and growth. As a statting point, consider a simple Robinson Crusoe economy of two people. Each devotes a percentage of their time to various activities and, if they have different talents, it pays for each to specialize: "you hunt, I farm, we trade."

Trade in this context faces several constraints. The first is what economists call the "double coincidence of wants." Only if both parties want what the other has, will trade occur. A second constraint is the transportation costs associated with bringing traders face to face. Unless the traders happen to be in close proximity to one another, each will have to incur travel costs to fmd a trading partner and to consummate the trade. Finally, there are

the costs associated with bilateral bargaining once traders do locate one another. Each trader's uncertainty about the other's value and natural desire to capture a larger share of the gains from trade means that each tries to gain the upper hand and this increases the likelihood of breakdown. For example, if one trader bargains using a take-it-or-leave-it strategy, there is the possibility that the other will leave it.

Cities reduce many of the costs associated with trade. First, the concentration of commerce at centralized locations, i.e. "markets," makes it easier to find someone who wants to trade. Second, by centralizing commerce in a single location, it is possible to locate near frequent trading partners to reduce travel costs. Third, and perhaps the biggest advantage that cities can offer, is that of efficient price formation. By reducing uncertainty about each item's value, prices reduce the bargaining costs associated with bilateral trade. In addition, prices send signals to buyers about what and how much to consume and signals to sellers about what and how much to produce. By coordinating buyer and seller activity, prices make economies more efficient, i.e. "wealthier."

Despite the advantages of cities as a way to organize commerce, it is easy to imagine a future in which further development of the Internet negates most of these advantages. First, the Internet makes it possible to transact remotely and asynchronously with a variety of trading partners. Certain kinds of transactions are better suited to the Internet than others, like those involving services. For example, banking services, particularly those associated with demand deposits, like checking or money market accounts, are already being delivered through on-line services. Other types of transactions, like those involving goods, as opposed to services, are less easily replaced. For goods with high transportation costs, it is difficult to foresee alternatives to local stores. It is difficult to imagine a virtual grocery store, for example. For goods that can be shipped long distances at relatively low cost, however, local stores will become obsolete. For example, the low transport costs of compact discs make it easy to forecast the demise of record stores or perhaps their evolution into related lines of commerce.

Lastly, it is not difficult to forecast that the price formation function of cities is going to be replaced by the Internet. Information about prices is already available through the Internet for certain types of goods, like computers. Advertising about prices, and private companies that disseminate price information over the Internet, in much the same way that Reuters or Bloomberg disseminate financial information, will largely replace the price formation function now provided by geographic markets in cities.

So, except for economizing on transportation costs, the Internet has the potential to negate most of the

economic advantages of cities as a way of organizing commerce. We refer to this as the "market widening" hypothesis.

Sherwin Rosen [1] modeled this phenomenon with respect to the development of television as a technological advance that lowered the costs of delivering entertainment to a wider audience. When remote sellers are brought into competition with one another, consumers benefit from the increased competition among sellers. The main prediction to emerge from his model is that small differences in the quality of goods or services offered by sellers lead to large differences in economic rewards, i.e. "skewness." Rosen labeled the better situated sellers "superstars" and pointed to the large profits earned by the best comedians, actors, singers, and athletes as evidence supporting his skewness hypothesis.

If we apply Rosen's model to the development of the Internet as a market widening technology, we can predict the disappearance of "local" markets for many types of goods and services--except those for which transportation costs are significant or those served by companies without significant economies of scale. Economies of scale are important because without them, it is impossible to serve a larger, or "wider," market.

We can also predict that sellers will react to the increased competition by trying to differentiate their products so they have something unique to offer; otherwise, they will be forced out of business from better situated competitors. The net effect is that consumers will be offered a wider array of consumption choices.

### 3. Firm Devolution

Up to this point, we have considered the effects of the Internet on the function of cities as markets. However, not all transactions occur through markets. Firms exist as an alternative to markets as a way of organizing economic activity. Certain types of transactions are less costly if consummated under the corporate umbrella. Arms length market transactions are replaced by a set of contracts between the firm and its employees. The contracts serve to coordinate economic activity and to align the incentives of the employees with those of the firm's owners. Monitoring by supervisors within the firm complements the incentive aligning function of the contracts.

Viewed this way, a geographically compact firm has many of the same advantages of a geographically compact market: it economizes on transportation costs and makes it easier to coordinate and monitor activities. And like geographically compact markets, it is not difficult to imagine a future in which further development

of the Internet improves communications to the point where many of the monitoring and coordination functions of a firm will be done remotely, or moved out from under the corporate umbrella into the marketplace. We call this "firm devolution."

Telecommuting, for example, moves transactions between firms and employees out from under the physical corporate umbrella. For telecommuting to be successful, remote monitoring and coordination must be complemented by incentive contracts that reward employees for output. The ultimate incentive aligning contract is to replace permanent employees with independent contractors, which would move them out from under the legal, as well as the physical corporate umbrella. The Internet makes it possible for telecommuters; or contractors to locate in remote, low cost areas. For example, many firms have already contracted out services like data entry, shipping, and phone answering to remote, independent contractors.

There are limits, however, on the degree to which firms can monitor remote employees or replace internal workers with independent contractors. Transportation costs, the difficulty of writing contracts and monitoring output, and the inability of some transactions to occur asynchronously place limits on the ability of firms to move transactions out from under the corporate umbrella or to remote locations.

#### 4. Testable Hypotheses

The market widening and firm devolution effects have similar implications. As geographical constraints tying consumers to markets and workers to firms are loosened, workers, consumers, and firms will move to more desirable and low cost locations. To the extent that geographical constraints are associated with urban living, one would expect a de-urbanization. There is disagreement, however, on the extent of this migration.

Gilder [2] believes that these changes will be the end of cities. Given a choice, people will move to an environment that has a higher quality of life. He sees the oncoming bandwidth revolution as allowing people to work anywhere they want. It will also allow people to enjoy the benefits of living in a city, like visiting a museum or going to the opera, virtually, in cyberspace, more easily than if they were actually there.

In contrast, Peters [2] suggests that cities will evolve into new entities that are only somewhat different than today's cities. He believes that people like the human contact associated with large cities. People like to "schmooze", and they like being physically near their competitors so that they can monitor their behavior.

While he acknowledges that agents and software could do this virtually, he does not believe there is any substitute for going to a ball park or having lunch with your competitor at a fine restaurant.

To "test" the migration hypothesis, it is important to understand the economic forces that limit the extent of this migration. For example, **let us assume that** San Francisco, California is more attractive than Lincoln, Nebraska. Citizens of Lincoln will move to San Francisco, and housing prices will be bid up in San Francisco and bid down in Lincoln. Eventually, the two cities will be equally attractive or Lincoln will become deserted. (e.g. Landsburgh [1]).

We will refer to the state in which all places are equally attractive as an "equilibrium." Now suppose that this initial equilibrium is perturbed by a technological innovation, like the Internet, that loosens the geographical constraints associated with work and consumption. As a consequence, until a new equilibrium is established, one would expect to see an exodus from certain cities and an immigration towards others.

It is probably too soon for such an exodus to show up in the data for several reasons. First the bandwidth is still too narrow to deliver many of the benefits associated with the Internet. Second, migrations take time. The technological changes are probably too recent to spur much migration. Lastly, the most recent population data are a year old, which is probably not recent enough to pick up many of the hypothesized changes.

Rather than focus on population migration, we construct a measure of Internet connectivity to test some of the same hypotheses. The connectivity measure should act as a "leading indicator" that is likely to pick up some of the population shifts before they occur. There are two reasons for this. First, workers who have chosen to live in attractive areas are the ones who have given up the most in terms of the benefits of geographic proximity to markets or work. They are the ones most eager to take advantage of the market widening or firm devolution effects of the Internet, and one would predict that they would be early adopters of the technology. Second, firms likely to take advantage of the Internet are more likely to be early movers to low cost or attractive areas. To the extent that Internet connectivity is an early indicator of future population shifts, cross sectional geographic variation in connectivity allows us to test some of these hypotheses about technology adoption.

#### 5. Data and Results

We use 3 digit zip codes as our basic unit of analysis. The use of zip codes has several advantages.

First, it is easier to match up data from different sources when using well defined zip code areas than by using less clearly defined "cities." This is especially true of the Internet connectivity measure which uses self reported data for location. Firms located in the environs of major cities have wide discretion of how to report their location. There is no such discretion for zip codes.

A second advantage of zip codes is that it is possible to use areas that more closely correspond to local geographic "markets" than do cities. Rand McNally [41] denotes a set of 3 digit zip codes that closely correspond to economic "trading" areas. This allows us to choose areas that more closely relate to the hypotheses of interest than would cities. We selected only those areas corresponding to economic areas, as defined by a Rand McNally "A" or "B" rating. This excluded areas like the Silicon Valley, where the 3 digit zip code does not closely correspond to the trading area. In addition, we selected only those areas with at least 200,000 population. This cut was made to reduce potential heteroskedasticity in the data. Adding small areas would potentially add outliers to the data.

The demographic data for this study was obtained from The Rand McNally Commercial Atlas & Marketine Guide f41. This source provided demographic data according to the 559 areas in the United States defined by 3-digit zip codes. Specifically, the Rand McNally data set included information regarding population, number of households, auto registrations, and retail sales for a given 3-digit zip code area.

In addition to data on demographics, this study relied on connectivity data. The connectivity data is the number of unique commercial Internet domains per 3 digit zip code as of June 20, 1995. The underlying data source is the InterNIC database, ' but it is sorted to throw out the duplicate domain names registered by corporations (compiled by Internet Info, Falls Church, VA). There is a trend toward registering domain names in order to stake out potentially valuable intellectual property. Simply counting the number of domains would overstate the actual number of companies. The weakness of the data is that it does not measure intensity of use by domain.

The final data source was a recent volume of Places Rated Almanac [5]. This almanac allowed the researchers to rank many of the zip areas with respect to quality of life. The actual number is the sum of ranks in 10 categories of the 343 largest areas, consequently the variable can take on any value from 10 to 3430. In the empirical work that follows, we treat the ranking as a cardinal variable.

## 5.1 Variable Description

The dependent variable, Internet connectivity, is defined as the log of Internet connectivity per capita for a zip area and is denoted by LOGCON. LOGCON was calculated by dividing the number of Internet domains per zip area by the area's population and taking the log of the quotient (Table 1).

Each zip area has a classification code associated with it. The classification code is intended to reflect the extent to which a zip area conforms to the actual trading area of the city. Zip areas classified as A conform closely to the actual trading area, and areas coded as B conform somewhat closely to the trading area. Areas coded as B1 are somewhat larger and areas coded as B2 are somewhat smaller than the actual trading area. The other classification codes, such as C and D, indicate zip areas that do not conform very closely to a specific trading area. These areas were not used.

The regressors for this study include demographic variables as well as a quality of life index and regional dummies. LOGPOP represents the log of a zip area's population; the population of an area is expressed in thousands of persons. Additionally, LOGSAL denotes the log of total sales in an area expressed in thousands of dollars. LOGSALUN represents the log of the number of Rand McNally Sales Units for a given zip area. LOGHOUSE and LOGAUTOS denote the log of the quantity of households and the log of the quantity of automobiles, respectively, for a zip area.

*Table 1 Variable Names and Definitions*

Definition

LOGPOP  
LOGSAL  
LOGSALUN  
LOGHOUSE  
LOGAUTOS  
LOGRANK  
REG[14]

LOGCON	log of (unique commercial domains per 1000 people) log of population log of sales log of Rand McNally Sales Units log of the number of households log of the number of automobiles log of city ranking mean of LOGCON for region 1 less the mean of region 4
REG[2-4]	mean of LOGCON for region 2 less the mean of region 4
REG[3-4]	mean of LOGCON for region 3 less the mean of region 4

The LOGRANK variable denotes the log of the cumulative ranking as outlined in the [Places Rated Almanac](#). The almanac ranks 343 metropolitan areas

according to a Borda count. For ten categories, such as housing, education, and climate, a ranking is assigned for each metropolitan area. For example, Abilene, Texas ranked 24th out of 343 metropolitan areas in the housing category. The rankings in the ten categories are added together to arrive at a cumulative ranking for each metropolitan area. The LOGRANK variable represents the log of this cumulative ranking.

Regional dummy variables were also used in this investigation. The United States was divided into four regions; the numbers 1,2,3, and 4 represent the East, South, Midwest, and West, respectively. The states that are included in each region are exhibited in Table 2. The variable REG[1-4] corresponds to the dummy variable for region one minus the dummy for region four, (EastWest). Similarly, REG[2-4] and REG[3-4] correspond to the dummy variables for (South-West) and (MidwestWest).

Table 2 The Four Regions of the United States

East	South
Connecticut	Alabama
Delaware	Arkansas
Maine	Florida
Maryland	Georgia
Massachusetts	Kentucky
New Hampshire	Louisiana
New Jersey	
New York	
Pennsylvania	
Rhode Island	
Vermont	
West Virginia	

- Mississippi
- North Carolina
- Oklahoma
- South Carolina
- Texas
- Tennessee
- Virginia

## 5.2 Comparing Connectivity

respectively, based on the size of the residuals. The domains per thousand measure is a raw measure of Internet connectivity. The residuals reported in Tables 3 and 4 are derived from a regression of LOGCON on demographics (regional dummies and attractiveness variables were not used).

In Table 3, the areas are ranked in descending order. Washington, DC is the most connected area. It is interesting to note that many of the well connected areas are in the Southwest. In addition, it appears as though proximity to government entities is important. The Air Force Academy is located in Colorado Springs, and Huntsville is the home of the Space Center. Three of the top ten are located in the areas within commuting distance of Silicon Valley.

The areas are arranged in ascending order in Table 4. Albany, GA is the least connected city. Interestingly, seven out of the ten least connected cities are located in the South. None of the least connected cities are located in the West or the East.

- Midwest
- Illinois
- Indiana
- Iowa
- Kansas
- Michigan
- Minnesota
- Missouri
- Nebraska

- West
- Alaska
- Arizona
- California
- Colorado
- Hawaii

Idaho  
 Montana  
 Nevada

*Table 3 The Ten Most Connected Areas*

Area  
 Washington, DC  
 Provo, UT  
 San Francisco, CA  
 Salinas, CA  
 Colorado Springs, C

North Dakota New Mexico Trenton, NJ  
 Ohio Oregon  
 South Dakota Utah  
 Wisconsin Washington  
 Wyoming

By regressing LOGCON on the demographic variables (excluding regional dummies and attractiveness), we are able to control for demographic factors like population and the number of households. This permits connectivity comparisons for areas with different demographics. We use the residual from the regression as a measure of connectivity. It measures how a region's connectivity compares to other regions with similar demographics. If a region has above average connectivity it will have a positive residual, and below average, a negative one.

Tables 3 and 4 list the commercial Internet domains per 1000 people and the residuals for the ten most connected areas and the ten least connected areas,

Austin, TX  
 Huntsville, AL  
 SantaRosa, CA  
 Portland, WA

Domains/100  
 0.746  
 0.28  
 2.197  
 0.323  
 0.728  
 0.412  
 0.914  
 0.267  
 0.335  
 0.378

Based on 170 3-digit zip code "trading areas"

*Table 4 The Ten Least Connected Areas*

Connectivity  
 2.716  
 2.383  
 2.212  
 2.177  
 2.086  
 1.944  
 1.933  
 1.904  
 1.623  
 1.467

Area	Domains/1000	Connectivity
Albany, GA	0.005	-2.365

Zanesville, OH	0.004	-1.889
St. Cloud, MN	0.032	-1.808
Tupelo, MS	0.004	-1.636
Rocky Mount, NC	0.015	-1.493
Macon, GA	0.014	-1.416
Lakeland, FL	0.02	-1.392
Shreveport, LA	0.013	-1.36
Jackson, TN	0.008	-1.356
Corpus Christi, TX	0.014	-1.244

Based on 170 3-digit zip code "trading areas"

### 5.3 Connectivity by Region and Attractiveness.

To test whether connectivity is related to different geographical regions, or the attractiveness of an area, we regressed our connectivity measure on a set of regional dummies and the area attractiveness variable, as well as the demographic controls. The results of the regression are reported in Table 5.

The coefficient on the attractiveness variable, LOGRANK, is statistically significant and negative. It is important to remember that the ranking variable is based on a Borda count. A low number corresponds to a relatively high rank, so that a negative sign on the LOGRANK parameter suggests that more attractive areas have a higher connectivity measure.

Other parameters of interest are the coefficients on the region dummies. We find that two of the\* variables are statistically significant, REG[2-4] and REG[3-4]. The regional dummy variables measure connectivity for regions 2 (South) and 3 (Midwest) relative to connectivity of region 4 (West). The West appears to be more connected than either the South or the Midwest. The other regional dummy, REG[1-4], is not statistically significant from zero. In words, the average connectivity for the East region is not significantly different from that of the West.

To summarize the results, the attractiveness of an area is related to its connectivity, and the West and the East are more connected than the South and the Midwest.

*Table 5 Regression Parameter Estimates*

Variable  
Intercept  
LOGPOP  
LOGSAL  
LOGSALUN  
LOGHOUSE  
**LOGAUTOS**  
**LOGRANK**  
**REG[ 1-4]**  
**REG[2-4]**  
**REG[3-4]**

## 6. Conclusions

Coeff/	Sid. Err.	t
2.799	3.223	0.87
-1.805	0.679	-2.66
2.23	0.649	3.43
-0.026	0.268	-0.1
0.366	0.6	0.61
-0.518	0.328	-1.58
-1.31	0.394	-3.33
0.198	0.111	1.78
-0.441	0.101	-4.38
-0.421	0.107	-3.93

R-SQUARE 0.573; n= 170

Our main findings suggest preliminary support for the hypothesis that Internet connectivity is related to

region and attractiveness. Since development of the Internet is a relatively recent phenomenon, we argue that the causality runs from attractiveness to connectivity. Specifically, cities that are highly ranked in terms of overall quality of life have a higher level of connectivity. Furthermore, there appear to be regional differences with respect to connectivity. The East and West are more connected on the whole as compared to the South and the Midwest.

Relating these empirical findings to the market widening and firm devolution effects of the Internet is a bit more difficult because one cannot rule out spurious alternative hypotheses. However, as preliminary results, they seem to support the hypothesis that firms, consumers, and workers will migrate to attractive areas as geographical constraints are loosened by the bandwidth revolution.

There are many avenues for further research on this topic. One of the first should be attempts to rule out alternative hypotheses for the correlation between connectivity and area attractiveness. Because the number of Internet domains is growing rapidly, one might analyze growth rates by region. As more detailed data become available, it may become possible to sort areas by the type of connectivity. This may make it possible to sort out the market widening from the firm devolution effects of the Internet. Additionally, one might determine how sensitive the results are to the type of ranking survey used.

## References

Prob>jtj  
0.387

1. Rosen, S., "The Economics of Superstars," The American Economic Review, Dec. 1981.

0.009  
0.001  
0.923  
0.543  
0.117  
0.001  
0.077  
0  
0

4.

2. Landsburgh, S. E., The Armchair Economist: The Economics of Everyday Experience, New York, NY: The Free Press, 1993.

3. Peters, T. and G. Gilder, "City vs. Country: Tom Peters & George Gilder Debate the Impact of Technology on Location," Forbes ASAP, Feb. 27, 1995.

Rand McNally 1995 Commercial Atlas & Marketing Guide, 126th Edition, Chicago, IL.: Rand McNally & Company, 1995.

5. Savageau, D. and R. Boyer, Places Rated Almanac: Your Guide to Finding the Best Places to Live in North America, New York, NY: Simon and Schir.ter, Inc., 1993.

6. Thomas, G. S., The Rating Guide to Life in America's Fifty States, Amherst, NY: Prometheus, 1994.