2007 V35 2: pp. 155–182
REAL ESTATE
ECONOMICS

Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems

Matthew E. Kahn*

Over 25 billion dollars were spent between 1970 and 2000 in 14 major cities in the United States on the construction of new rail transit lines. This massive investment in rail transit construction and expansion allows me to study the consequences of local public goods improvements for communities nearby new stations. This article uses a 14-city census tract–level panel data set covering the years 1970 to 2000 to document significant heterogeneity in the effects of rail transit expansions across the 14 cities. Communities receiving increased access to new "Walk and Ride" stations experience greater gentrification than communities that are now close to new "Park and Ride" stations.

Over 25 billion dollars were spent between 1970 and 2000 in many major cities in the United States on the construction of new rail transit lines. Billions more have been spent on maintaining and improving existing rail transit lines. While the supply of rail transit has increased, the fraction of metropolitan-area workers commuting using public transit has declined from 12% in 1970 to 6% in 2000 (Baum-Snow and Kahn 2005).[1]

Rising incomes and suburbanization of jobs and people explain a large fraction of this decline (Glaeser and Kahn 2004). In 1990 in Boston, 36% of workers who lived and worked in the center city commuted on public transit, compared to 5% of workers who lived and worked in the suburbs.[2]

*Institute of the Environment, University of California Los Angeles, Los Angeles, CA 90095 or mkahn@ioe.ucla.edu.

[1] "So why has transit enjoyed such political success in recent decades? It appeals to interests across the political spectrum: downtown and construction-related businesses, construction and transit labor unions, environmentalists, good-government organizations, advocates for the poor, and a wide variety of others who perceive transit as a way of reconciling development, equity and amenity goals" (Altshuler and Luberoff 2003, p. 217).

[2] Using year 2000 Census of Population and Housing data for all census tracts within 20 miles of a Central Business District (CBD), I estimated the following regression: *Percent Using Public Transit* $= 0.288 - 0.042 * \log($*Tract Mean per-capita income*$)$

Proponents of rail transit investment argue that this transit mode promotes environmental sustainability and helps to strengthen center city economic viability. I define transport-oriented communities as places that experience increased rail transit access, especially those locations that are now close to "Walk and Ride" stations due to investments in extensions of existing rail lines and creation of new lines. The supply of such locations has increased over time. This article examines the consequences of transit expansions in 14 of the 16 major cities that expanded and built rail transit systems between 1970 and 2000.[3] Across the 14 major cities studied in this article, 1,351 census tracts that were far from a rail transit station in 1970 are within one mile of a rail transit station in the year 2000.

This massive investment in rail transit construction and expansion allows me to study the consequences of local public goods improvements for communities nearby new stations. Using geographic information systems (GIS) software, I calculate each tract's centroid's distance to the closest rail transit station in each census year. Transit expansions reduce a census tract's distance to the closest station. My census tract level panel data set covers 14 major metropolitan areas over the years 1970 to 2000.

I first examine what observable attributes of communities predict being "treated" with increased access to rail transit. Such descriptive facts are useful for considering the political economy of who demands increased access to transit. In the second half of the empirical work, I examine how key community outcomes, such as average home prices and the share of the community who are college graduates change in *treated* tracts versus in *control* tracts. My control group is census tracts in the same metropolitan area with similar observable attributes that have not experienced increased proximity to rail transit.

By documenting how large transportation infrastructure investments affect population sorting and home price capitalization within multiple cities, this article builds on two real estate literatures. The first literature examines rail transit access capitalization into real estate prices (for a study of Chicago see McMillen and McDonald (2004) and for a study of London see Gibbons and Machin (2005)). This article also contributes to the literature estimating the demand for living in "new urban" communities (see Tu and Eppli 1999 and Coulson and Lahr 2005).

$+ 0.031 * \log(\textit{Tract population density})$. The mean of the dependent variable equals 0.083. In this census tract–level regression, metro-area fixed effects are included and there are 41,301 observations.

[3] Due to data limitations, I do not include Buffalo or St. Louis in this study.

I document evidence of heterogeneous treatment effects both within cities and across cities. Communities treated with a new "Walk and Ride" station are more likely to gentrify than communities treated with a new "Park and Ride" station within the same metropolitan area.[4] Across the 14-city sample, two cities that stand out in terms of overall gentrification effects are Boston and Washington, DC. Public transit expansions may not gentrify every community they reach. Glaeser, Kahn and Rappaport (2007) argue that public transit stations often act as a poverty magnet. The urban poor are less likely to own cars and thus place a greater value on rail access. In some metropolitan areas such as Atlanta and Los Angeles, for example, I find that the share of college graduates living in communities near new "Park and Ride" stations declines relative to trends in control tracts. This could partially explain why so many rich suburban communities (such as Georgetown in Washington, DC) fear the extension of public transit into their community (Bowes and Ihlanfeldt 2001, Altshuler and Luberoff 2003).

Data

For this analysis, I use home price and demographic data at the census tract level, and historical data on the location of rail transit lines and railroads. Throughout this article, I am mainly interested in how communities close to new rail transit stations evolve once "treated" relative to time trends in "control" communities. Given my spatial focus on community dynamics, census tract panel data is more useful than individual real estate sales.[5]

The census tract data is from the Urban Institute and Census Geolytics' Neighborhood Change Database. This is a set of repeated cross sections from the 1970, 1980, 1990 and 2000 decennial censuses at the census-tract level normalized to

[4] This article is motivated by one specific example of gentrification near a new "Walk and Ride" rail transit stop. The Davis Square area is located 4.4 miles from Boston's CBD and 1.5 miles from Harvard University. In 1970, the closest rail transit stop was Harvard Square. In 1970, 8.2% of Davis Square's adults were college graduates. In 1985, the Red Line was opened in Davis Square. Residents of this community now had easy access to a "Walk and Ride" station offering a fast subway ride to Harvard Square and downtown Boston. In the year 2000, 49.7% of Davis Square's adult population was college graduates. For the Boston metro area, 14.7% of adults were college graduates in 1970 and 39.6% in the year 2000.

[5] Individual real estate sales would offer more precise home price information and better structure control variables but collection of such micro data with precise geographical information on the exact street address for 14 cities in 1970, 1980, 1990 and 2000 would be very costly. In addition, such micro transaction data would represent the select set of recently sold units and would provide no demographic data on the attributes of the residents. In contrast, census data represents a representative sample of who lives in the community. It is true that census home price data are self reported. Throughout this article I am interested in both demographic dynamics in treated communities and home price dynamics in such communities.

2000 tract geography.[6] These data contain the evolution of demographic characteristics and transit ridership for exactly the same geographic areas over time. Census tracts are sufficiently small geographically such that they capture much of the demographic variation across neighborhoods.[7]

Using GIS software, I observe each census tract's distance to the CBD, its distance to the closest rail transit station and its distance to the nearest railroad line (see Baum-Snow and Kahn 2005). CBD definitions are taken from the 1982 Economic Censuses Geographic Reference Manual. They represent agglomerations of census tracts that surveyed local business leaders reported to represent the center of economic activity for each metropolitan region. The basis of the rail transit data is the Bureau of Transportation Statistics' 2003 National Transportation Atlas Database (NTAD). The NTAD includes digital maps of rail transit lines and stations for most U.S. cities.[8] I only include modern rapid transit lines, not vintage trolleys or commuter rail lines. For more details on the data set construction see Baum-Snow and Kahn (2005). The NTAD also provides digital maps of railroads from the mid-1990s. I use GIS software to construct the distance from each census tract to the nearest railroad lines.

The full set of transit lines in the data set is detailed in Table 1.[9] This table highlights that cities differ with respect to the extent of the treatment's size and years until completion. For example, the majority of San Francisco's rail transit system was built in the early 1970s, but the system was expanded in the mid-1990s. In Washington, DC, there was new construction of its rail transit

[6] A full explanation of the tract conversion process and the potential errors that might be raised is available through Harvard's Graduate School of Design's Web site where "Special Issues in Normalizing Multi-Year Census Data" are analyzed for the Neighborhood Change Database at http://www.gsd.harvard.edu/gis/manual/censuscd/ncdb_docs/SpecialIssues.pdf.

[7] Earlier research has examined how rail transit access affects nearby land use patterns. Cervero and Landis (1997) focus solely on San Francisco, and their geographical unit of analysis is a "superdistrict." The San Francisco metropolitan area is partitioned into 34 of these districts. With this level of spatial aggregation it is difficult to study how "nearby" communities are affected by this treatment.

[8] In the areas for which NTAD data is not available or up to date, I constructed digital maps of lines and stations based on digital street maps and physical maps of transit lines' locations to reflect infrastructure as of January 1, 2004. Baum-Snow and I use transit construction histories from various sources to form digital maps of the transit infrastructure and stations on January 1, 1970, 1980, 1990, 1994 and 2000 in addition to 2004.

[9] The Web site world.nycsubway.org contains informative details and photos of the trains used in each rail transit city.

Table 1 ■ Rail transit construction 1970 to 2000.

City	Line	Open By	Cost Per Mile (Millions)	Type of Construction
Los Angeles	Red Line	Jan 1993	330	T
	Red Line	July 1996	245	T
	Red Line	2000	227	T
	Blue Line	1990	40	R
	Green Line	1995	36	H/E
	Gold Line	July 2003	63	R
Sacramento	North/East Line	1987	10	S/H/R
	East Line	Jan 1998	15	S/R
	South Line	Dec 2003	35	R
	East Line	June 2004	32	S/R
San Diego	Blue Line	July 1981	8	R/S
	Orange Line	March 1986	7	R
	Orange Line	1989	10	R
	Orange Line	1990	32	R
	Blue Line	1992	?	R
	Blue Line	1996	37	R
	Blue Line	1997	37	E
	Orange Line	Sep 1998	30	S
San Francisco	BART	Sep 1972	16	T/E/R
	BART	Jan 1973	16	T/R
	BART	May 1973	16	H/R
	BART	Nov 1973	16	T
	BART	Feb 1996	106	T
	BART	Dec 1996	?	?
	BART	May 1997	?	H
	BART	June 2003	106	T
	MUNI	1998	37	S
San Jose	VTA	June 1988	25	S
	VTA	Aug 1990	25	H
	VTA	April 1991	25	H
	VTA	Dec 1999	42	S
	VTA	May 2001	39	S/E
	VTA	June 2004	54	S
Denver	D	Oct 1994	21	S
	C/D	July 2000	22	S
	C	April 2002	24	R
Washington	Red Line	Jan 1977	?	T/E
	Blue Line	July 1977	82	T/H
	Red Line	Feb 1978	?	R
	Orange Line	Nov 1978	?	R
	Orange Line	Dec 1979	106	T
	Blue Line	Nov 1980	?	T
	Red Line	Dec 1981	?	T
	Yellow Line	Apr 1983	?	T

Table 1 ■ (Continued)

City	Line	Open By	Cost Per Mile (Millions)	Type of Construction
	Blue Line	Dec 1983	?	R
	Red Line	1984	?	R
	Orange Line	Jun 1986	26	H
	Red Line	Sep 1990	?	T
	Green Line	May 1991	?	T
	Blue Line	Jun 1991	?	R
	Green Line	Dec 1991	?	T
	Green Line	Dec 1993	?	R
	Blue Line	Jun 1997	53	R
	Red Line	Jul 1998	162	T
	Green Line	Sep 1999	222	T
	Green Line	Jan 2001	138	T/H/E
Miami	Metrorail	1985	48	R/E
	Metrorail	May 2003	?	E
Atlanta	East/West	1979	33	R/T/E
	North/South	Dec 1981	33	T
	North/South	Sep 1982	33	T
	North/South	Dec 1984	33	T
	North/South	Aug 1986	33	R
	North/South	Dec 1987	33	R
	North/South	Jun 1988	33	R
	East/West	Dec 1992	33	E
	North/South	Dec 1992	33	R
	East/West	Jun 1993	33	R
	North/South	Jun 1996	52	H
	North/South	Dec 2000	33	E
Chicago	Blue Line	Feb 1970		S/H
	Blue Line	1984	?	H
	Orange Line	1993	56	R
	Green Line	1994	?	E
Baltimore	Metro Subway	1983	100	T/E
	Metro Subway	1987	?	H
	Metro Subway	1994	?	T
	Light Rail	1993	18	S/H/R
	Light Rail	1997	14	R/N
Boston	Orange Line	1975	?	E
	Orange Line	1975	?	R
	Orange Line	1977	?	R
	Red Line	Sep 1971	?	R
	Red Line	Mar 1980	?	R
	Red Line	Mar 1985	179	T
	Orange Line	May 1987		E
	Orange Line	May 1987		R
	Green Line	1985		S

Table 1 ■ (Continued)

City	Line	Open By	Cost Per Mile (Millions)	Type of Construction
Portland	MAX	1986	14	R/S
	MAX	Sep 1998	54	S/T/H
	MAX	July 2001	23	E/H
	MAX	May 2004	63	S
Dallas	DART	May 1997	43	S/R/T
	DART	Dec 2002	43	R/H

Note: Construction types codes are as follows: R = Railway Right of Way, T = Tunnel, S = Street, H = Highway Median and E = Elevated.

system in the 1970s, 1980s and 1990s. The table also provides some data on the cost per mile by project. There are huge differences in these nominal costs across cities. San Diego's Blue Line cost $37 million per mile, while part of Los Angeles' Red Line cost $330 million per mile.

For each of the 14 rail transit systems, I accessed their main Web page and collected information on whether each station stop had a parking lot. This information allows me to partition all station stops into "Walk and Ride" versus "Park and Ride" stations. Below, I discuss in detail why these different types of stations are likely to be associated with different outcomes for treated communities.[10]

Table 2 documents the count of census tracts by metropolitan area that were "treated" with increased access to rail transit due to rail transit extension and construction. Throughout this study, I focus on the sample of census tracts that are within 20 miles of the 14 CBDs listed in Table 2. In addition, I only include those tracts that were more than a mile from rail transit in 1970. Across the 14 cities, by the year 2000, 816 census tracts that were more than one mile from a rail transit station were now within one mile of a "Park and Ride" station, and 559 census tracts that were more than one mile from a rail transit station were now within one mile of a "Walk and Ride" station. Given that the average census tract contains roughly 4,000 people, a large number of people live close to new rail stations. Table 2 also highlights that, across the 14 cities, only 18% of the census tracts that were more than a mile from rail transit in 1970 were treated.

[10] The largest parking lots are often built at the last stop of a line. The Boston Alewife parking lot on the Red Line has 2,600 spots.

Table 2 ■ Tract count by treatment status.

	Total Tracts	New "Park and Ride" Tracts	New "Walk and Ride" Tracts
All 14 Cities	7590	816	559
Atlanta	426	53	30
Baltimore	518	85	33
Boston	414	27	14
Chicago	666	52	32
Dallas	549	23	32
Denver	502	10	11
Los Angeles	1567	79	159
Miami	326	42	10
Portland	373	23	53
Sacramento	287	34	12
San Diego	434	73	36
San Francisco	473	107	26
San Jose	328	36	27
Washington, D.C.	727	172	84

Note: The sample consists of all census tracts located within 20 miles of the 14 CBDs listed in this table that were more than one mile from the closest rail transit station in 1970. The count of "at risk" tracts is presented in the left column. The middle and right columns of this table report the count of tracts in each city that, due to rail transit expansions, are within one mile of a rail transit station in 2000 but were not within one mile of a rail transit station in 1970.

Rail Transit Supply

In most metropolitan areas considering new transit projects, a regional planning organization devises a detailed plan that forms the basis of a funding request from the Federal Transit Administration. The planning processes used by these regional planning organizations are not systematic across metropolitan areas. However, there are some transit line location patterns that persist across metropolitan areas. For example, all rail transit systems in the United States are oriented to serve the CBD. In Washington, DC, "the network was designed with a spoke-hub distribution paradigm, which makes the subway ideal for getting from a suburb to any part of the city, or vice versa, but unattractive for suburb-to-suburb travel; a Purple Line has been occasionally proposed to remedy this."[11]

Federal funding typically covers between 50% and 75% of the cost of new rail transit construction. The Federal Transit Administration assigns funding under its "New Starts" program based on evaluation of proposals from local

[11] See http://www.answers.com/topic/washington-metro.

governments looking to build new transit lines. Every year it draws up a list of priority projects that get funded according to the amount of money allocated by Congress. There are a few lines built exclusively with local funding.

The Political Economy of Which Communities are Selected to Receive Increased Rail Access

From a representative community member's perspective, rail transit access offers a trade-off (Bowes and Ihlanfeldt 2001). The rail line increases the set of possible commuting choices. Depending on what type of transit is built, it could bring about gentrification, but it may lead to more crime in one's town as the urban poor have greater access to the community. Introducing a "Park and Ride" station could lower local quality of life because of increased noise, traffic and congestion. Community homeowners have an incentive to lobby for the outcome that raises their home prices. If new transit is expected to lower local home prices, then the community will lobby against having a new rail transit station.

Lobbying can change the implementation of rail transit plans.[12] In the case of the siting of the Los Angeles Red Line, Altshuler and Luberoff (2003) stress that the transportation agency followed the path of least resistance:

> As originally conceived, this segment was to run for about five miles under Wilshire Boulevard, the most densely developed corridor west of the Mississippi River, before turning north toward the San Fernando Valley. In the mid-1980s, however, homeowners in a mid-Wilshire neighborhood objected that the line would stimulate too much development in their vicinity and might expose them to more crime by facilitating travel from low-income areas . . . When Congress did eventually approve funding for the line, it did so with a proviso (inserted by Representative Henry Waxman who represented mid-Wilshire) that it could not pass through the mid-Wilshire area. Instead, it would now turn north before it reached the neighborhoods in which substantial opposition had emerged and it would serve Universal City area. Because the new alignment was through less dense areas, it had considerably lower patronage potential than the original routing, but it faced no significant opposition—and that was what counted politically. In a similar vein, localities frequently minimized controversy by siting new transit lines in existing rail or freeway corridors, chosen for their availability rather than

[12] One underexplored determinant of spatial differences in investment in rail is which political party controls the Mayor's and Governor's Office. A political bias often exists such that Democrats are more likely to invest in urban stations while Republicans are more likely to build suburban "Park and Ride" stations. I thank Terry Regan for this point.

> their optimality from a patronage standpoint. This strategy was used, for example, in Portland and Sacramento and for portions of the Los Angeles' new Blue Line (Altshuler and Luberoff 2003, p. 204).

This example highlights that, when communities view public transit as bringing in people whose income are below the average incumbents, they are more likely to oppose transit expansion.

The decision over whether to construct a "Park and Ride" station depends on a variety of supply and demand factors. Demand factors include how many potential park and riders live in neighboring communities. Supply factors include land availability and road access. For many built-up urban areas, there is no place to site a parking lot. In Boston, nine criteria are used to rank possible locations when deciding whether to build a "Park and Ride" station. These include: (1) access to roads, (2) available land, (3) demand model forecasts, (4) capacity, (5) cost per space, (6) environmental issues such as remediation of contaminated land, (7) implementation, (8) community support and (9) funding options. Transit Authorities prefer to not build parking lots because they lose money on these investments. They are costly to build (averaging up to $30,000 per spot to construct) and, unlike a for-profit garage, their rates are capped.[13]

Characteristics of "Treated" and "Control" Tracts

The possible treatments are no increase in access to rail transit, a "Walk and Ride" station built close to the community or a "Park and Ride" station. In the case of a "Walk and Ride" treatment, communities close to the proposed new station are likely to have a voice in whether they support this change. In the case of a "Park and Ride" treatment, the "community" is broader and will include all census tracts in a wider driving radius around the proposed new parking station. A parking station is more likely to be built if land is available and demand models predict that within a radius around the station there is wide interest in commuting using this new mode.[14]

Based on observable demographics such as income and educational attainment, how do census tracts that never are close to a rail transit station compare to tracts

---

[13] Communities are aware that there are negative externalities associated with "Park and Ride" stations and try to site these at town borders to "externalize" these costs on neighboring towns.

[14] A "Park and Ride" treatment can lower quality of life in the immediate surrounding area but increase quality of life for potential commuters who live a few miles away who can drive a short distance and "Park and Ride." In such a case, the immediate community may not have the political clout to block the project.

Table 3 ■ Community demographic attributes by treatment type.

	All	Treated with Park and Ride	Treated with Walk and Ride	Never Near Transit
1970 Census Data				
Average household income	54596.400	45458.400	47838.800	56489.560
Average home price	96364.140	77219.580	76403.640	100688.700
Population per square mile	8620.916	12300.920	14518.620	7575.992
% of Adults who are college graduates	0.140	0.111	0.137	0.145
% of Population living in poverty	0.112	0.171	0.213	0.095
% Black	0.169	0.378	0.259	0.124
% Hispanic	0.094	0.089	0.123	0.092
Miles from the CBD	9.219	6.397	4.121	10.071
% of workers commuting using public transit	0.109	0.185	0.212	0.088
Count of tracts	7647	816	559	6274

Note: This table reports sample averages for four different sets of census tracts based on 1970 data. The sample includes all tracts located within 20 miles of one of the 14 metropolitan areas listed in Table 2 that were more than one mile away from the closest rail transit station in 1970. *Treated with Park and Ride* are the subset of tracts that were more than one mile from a "Park and Ride" station in 1970 and are within one mile of a "Park and Ride" station in the year 2000. *Treated with Walk and Ride* are the subset of tracts that were more than one mile from a "Walk and Ride" station in 1970 and are within one mile of a "Walk and Ride" station in the year 2000. *Never Near Transit* are the subset of tracts that have always been more than one mile from the nearest rail station stop. Prices are measured in 1999 dollars. CBD = central business district.

that are "treated" with close access to a rail station? In Table 3, I report four sets of sample averages based on 1970 demographic data. Table 3 shows that the average tract that is treated with either a "Walk and Ride" station or a "Park and Ride" station is poorer, has higher population density, has higher poverty rates, has higher public transit use and is closer to the CBD than the average tract that has never been within one mile of a rail station stop. The average community treated with a "Park and Ride" station has a larger 1970 black population share than other communities. Average household income and home prices are roughly equal in treated "Park and Ride" and "Walk and Ride" communities.

To further explore the determinants of treatment as a function of observable tract attributes, I estimate linear probability models. The dependent variable is a dummy variable that equals one if a tract is within one mile of rail transit in the

Table 4 ■ Determinants of transit treatment for tracts not close to rail transit in 1970.

	All		All		Walk and Ride	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*
% Using Public Transit in 1970			0.414	0.059	0.098	0.042
log(Average household income in 1970)	−0.125	0.013	−0.093	0.014	−0.037	0.010
log(Population density in 1970)	0.006	0.004	0.003	0.004	−0.011	0.003
log(Distance from CBD in meters)	−0.204	0.007	−0.168	0.009	−0.163	0.006
log(Distance from railroad line in meters)	−0.043	0.003	−0.044	0.003	0.006	0.002
Constant	3.705	0.145	3.019	0.174	2.055	0.124
Mean of dependent variable	0.183		0.183		0.074	
Metropolitan area fixed effects	Yes		Yes		Yes	
Observations	7458		7458		7458	
Adjusted R squared	0.273		0.278		0.202	

Note: The sample includes the set of census tracts that were more than one mile away from a rail transit stop in 1970. This table reports three linear probability models. The dependent variable is a dummy that equals one if a tract is within one mile of transit in the year 2000. The sample includes tracts within 20 miles of the central business district (CBD).

year 2000. Recall that this sample only includes tracts that were more than one mile from rail transit in 1970. In Table 4, I report three linear probability models to examine how baseline community demographics in 1970 are correlated with subsequent treatment. The left column of Table 4 shows that poorer communities that are closer to the CBD are more likely to be treated. All else equal, proximity to a railroad line increases the probability of being treated. In the middle column of Table 4, I include a census tract's 1970 share of workers who commute using public transit. All else equal, communities where more people commute by public transit were more likely to be treated with rail transit access. A 10 percentage point increase in the share of workers commuting by public transit increases the probability of being treated by 4.1 percentage points. This is suggestive evidence that communities that use public transit have lobbied for increased access to a faster public transit mode (*i.e.*, rail transit). In the right column of Table 4, the dependent variable is a dummy that equals one if the census tract was treated with access to a new "Walk and Ride" station between 1970 and 2000. Recall that access is defined to mean that the tract is within one mile of the closest station. The most interesting result reported in this column is the reduction in the income effect. Increases in tract income have a smaller effect on reducing the probability of a "Walk and Ride" treatment relative to the total effect reported in the left and middle columns of Table 4.

In the results section, I will present panel estimates of how increased access to "Walk and Ride" and "Park and Ride" stations affect community home prices and the share of college graduates who live in these communities. The interpretation of such treatment effect regressions hinges on whether communities differ with respect to how such treatment will affect them (heterogeneous treatment effects) and whether *ex ante* each community is aware of their own treatment effect. Put simply, new rail transit may offer heterogeneous treatment effects, and these differences in expected benefits are not observed by the econometrician. Those communities that expect to gain the most are more likely to seek treatment. This raises selection issues when I estimate outcome equations below (see Moffitt 1999). I am likely to only observe treatment in those communities that expect to gain from treatment. Below, I will discuss how selection on unobservables affects the interpretation of the results and present some instrumental variables estimates. The instrumental variable will be a tract's 1970 distance from the nearest railroad line. As shown in Table 4 (which can be thought of as a first-stage regression in two-stage least squares), tracts that are further from railroad lines are less likely to be treated. Proximity to a railroad line acts as a cost shifter. It is cheaper to convert old railroad lines into rail transit lines than to build entirely new rail infrastructure.

Empirical Framework for Estimating Community Treatment Effects

Using the census tract panel data set, I now report estimates of how key outcome measures, such as average home prices and the share of adults who are college graduates, evolve in "treated" versus "control" tracts. Throughout this section when I use the term "treatment effect" I mean that a census tract that was more than one mile from the nearest rail transit station is now within one mile of a rail transit station due to rail transit construction and expansion.

Using the 1970 to 2000 census tract panel, I observe each tract at four points in time. Define m as the metro area and define $Close_{jt} = 1$ as a dummy variable that equals one if tract j is within one mile of a rail transit station in year t.[15] The outcome regression equation is presented in Equation (1).

---

[15] It is certainly true that a household living more than a mile away from a "Park and Ride" station can benefit from the creation of such a station. In this article I am interested in how communities that have recently become connected to rail transit evolve. My implicit assumption here is that a one-mile ball around the station is sufficient for defining the station's community.

$$Y_{jmt} = \Phi_{mj} + \Phi_{mt} + B_{m1}Close_{jt} + B_{m2}\,Close\ to\ Walk\ \&\ Ride_{jt} \\ + B_{m3}\,Close_{jt} * X_{jm1970} + B_{m4}Year\ Dummy_t * Distance\ to\ CBD_j \\ + U_{jmt}. \quad (1)$$

In this regression equation, the unit of analysis is tract j in metro area m in year t. I control for tract and year fixed effects (Φ_{mj}, Φ_{mt}).[16] Estimates of B_{m1} represent a "double difference" between 10-year first differences in Y_{jmt} for treated tracts and the 10-year first difference in Y_{jmt} for observationally identical control tracts.

What is the treatment? As I discussed in the previous section, increased proximity to rail transit represents a bundle of direct and indirect effects. The direct effect is increased commuting options. The indirect effects depend on whether the marginal migrants to treated communities are richer or poorer than the incumbents.

I seek to test for heterogeneity within a city along several dimensions. Communities that gain increased access to "Walk and Ride" stations may experience different effects than communities that gain access to a "Park and Ride" station. The "Walk and Ride" communities are more likely to offer a New Urbanist lifestyle, while the "Park and Ride" communities may feature higher usage and more traffic congestion. Estimates of B_{m2} allow me to test this hypothesis. A second dimension of heterogeneity is captured by estimates of B_{m3}. This coefficient allows me to test whether the consequences of a tract's treatment status depends on its initial household income. To capture such heterogeneity while keeping the interpretation of the interaction terms manageable, I simply partition metro-area tracts based on their 1970 per-capita income. I create a dummy variable that equals one if the tract's 1970 per-capita income was above the metropolitan area's median. In the regressions I present below, this dummy is interacted with the *Close* dummy. I estimate Equation (1) separately for each of the 14 cities.

In estimating Equation (1), using a linear panel estimator, I am assuming that the error term is uncorrelated with the explanatory variables.[17] With tract, year

[16] In previous work (Baum-Snow and Kahn 2005), I documented the importance of controlling for metropolitan statistical area (MSA)-level trends that vary within the metro area. For example, if employment is suburbanizing over time, then census tracts at the fringe may offer shorter commutes in the year 2000 relative to when more jobs were downtown in 1970. The set of variables $Year\ Dummy * Distance\ to\ CBD_j$ allows me to control for such patterns. Intuitively, my control group is untreated tracts within the same metro area that are equidistant to the CBD as tracts that have been treated with rail transit access.

[17] I am also assuming that, conditional on tract fixed effects, U is independent and

and distance from the CBD multiplied by year fixed effects, this assumption would be invalid if there is an omitted variable that varies across census tracts over time that is correlated with treatment status. For example, suppose that for exogenous reasons crime disproportionately declined in the tracts treated with a new "Walk and Ride" station. The econometrician does not see this but the households would. In this case, I would observe gentrification and rising home prices and would attribute this to the rail treatment when it really reflects the value of an omitted variable (increased local safety).

Equation (1) imposes that, holding X_{jm1970} constant, the transit treatment effect is homogeneous within a city. In this equation, I did not subscript the regression coefficients with a tract-specific effect. If $B_{jm} = B_m$ (treatment homogeneity) and if $E(U_{jmt} \mid X_{jmt}) = 0$, then my estimates of Equation (1) yield consistent estimates of the treatment effects. But, suppose that communities do differ with respect to their treatments, and suppose that each community knows its B_{jm}.[18] In this case, a selection on unobservables issue will arise. Borrowing from the heterogeneous treatment effects literature (Heckman and Vytlacil 1999, Moffitt 1999), if communities differ with respect to what they expect to gain from receiving rail access, then those communities that gain the most will be more likely to be treated. Recognizing the possibility that the treated communities represent a select sample of those that would gain the most from treatment, I contrast my panel estimates with estimates based on an instrumental variables strategy.[19]

Results

Home Price Dynamics

Home prices might rise in treated communities due to the direct effect that residents of such communities now have a lower cost of traveling to the city center. New "Walk and Ride" stations would also guarantee increased foot traffic in

---

identically distributed (i.i.d.). This implicitly assumes that time-varying unobservables do not spill-over from one tract to adjacent tracts.

[18] If each community does not know its treatment effect, then the existence of heterogeneous treatment effects will not induce selection on unobservables (Heckman, Urzua and Vylacil 2006). When there are heterogeneous returns to treatment and the decision makers are choosing treatment status based on their knowledge of these returns, then "treatment on the treated" estimates will overstate the true average treatment effect (Moffit 1999).

[19] In the presence of heterogeneous treatment effects, there are really two separate research questions. My study's estimates provide new insights about the effect of "treatment on the treated." This is a different question than asking: "If a random census tract were treated with increased access to rail transit, how would its demographics evolve?" (see Moffit 1999).

the local community as people walked to the station. This new urbanist lifestyle might attract upper income households who are richer than the average incumbent. If the rich have a higher marginal utility for living in richer communities, this multiplier effect will feed on itself (Bayer, Ferreira and McMillan 2005). This gentrification will lead to improvements in retail stores and restaurants. One force going against the gentrification trend is the quality of the housing stock. If the rich have a strong taste for new, large housing and if the communities treated with a "Walk and Ride" station have old, small housing then the rich will be less likely to move in (Bond and Coulson 1989). I predict that there will be less gentrification in communities very close to a new "Park and Ride" station. The construction of such a parking lot creates a "lose/lose" for wealthy incumbent members of the tract. They do not want to use this rail transit mode and the quality of life in their community will fall due to congestion, traffic and rising crime exposure.

To test these hypotheses, I estimate Equation (1). I trim the home price distribution by only keeping tracts in each census year whose average home price is between the 5th and 95th percentiles of a metropolitan area's empirical distribution.

Table 5 reports 15 estimates of Equation (1). In the upper panel, I report the pooled estimate across the 14 cities and then each subsequent entry of the table reports a city-specific estimate of Equation (1). Across the 14 cities, home prices in tracts whose income was below their metro-area median in 1970 experienced a 3% increase in home prices when treated with a "Walk and Ride" station and no change in home prices in tracts treated with a "Park and Ride." Initially, above-median income communities treated with either a "Park and Ride" or a "Walk and Ride" experienced a small reduction in home prices relative to control tracts.[20] This average effect masks large differences across cities. As shown in Table 5, "Walk and Ride" stations have a positive and statistically significant effect on home prices in 6 of the 14 cities, while "Park and Ride" station treatment has a negative and statistically significant effect on home prices in 2 of the 14 cities (Portland and San Francisco).[21] Perhaps surprisingly, the estimates B_{m3} (the interaction term on tract transit treatment multiplied by the dummy for the tract having a 1970 average income greater than the metro-area median) does not tell a consistent story. In some cities such as Boston and Washington, DC, it equals zero, while in Sacramento, San Diego and San

[20] Recall that the control group is census tracts in the same metropolitan area that are equidistant to the CBD and have never been within a mile of a rail transit station.

[21] Note that Equation (1) imposes that the adjustment process is completed by the next census year. Below, I will test for adjustment lags by examining long differences from 1970 to 2000 for treated and control tracts.

Table 5 ■ Home price dynamics in treated census tracts.

Dependent Variable is the log of tract average home prices	All 14 MSAs		Atlanta		Baltimore		Boston	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	0.001	0.010	0.110	0.048	0.065	0.027	−0.050	0.034
Close to walk and ride station	0.026	0.014	0.219	0.080	0.167	0.039	0.120	0.070
Close to transit station∗ract income above msa 1970 median	−0.031	0.013	0.013	0.065	−0.203	0.036	−0.012	0.071
Observations	26514		1498		1832		1487	
R squared	0.94		0.878		0.951		0.964	
p value on *F* test for joint significance of the three variables	0.011		0		0		0.218	
	Chicago		Dallas		Denver		Los Angeles	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	0.106	0.042	0.050	0.048	0.333	0.059	0.018	0.026
Close to walk and ride station	0.103	0.053	0.069	0.067	−0.104	0.138	0.035	0.033
Close to transit station*tract income above msa 1970 median	−0.044	0.052	0.187	0.124			−0.042	0.032
Observations	2368		1878		1724		5521	
R squared	0.947		0.911		0.904		0.943	
F test for joint significance of the three variables	0		0.016		0		0.21	

(Continued)

Table 5 ■ (Continued)

	Miami		Portland		Sacramento		San Diego	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	−0.008	0.049	−0.112	0.042	0.176	0.042	−0.004	0.026
Close to walk and ride station	−0.125	0.128	0.089	0.039	0.090	0.049	−0.016	0.039
Close to transit station*tract income above msa 1970 median	0.146	0.091	−0.026	0.037	−0.294	0.052	−0.063	0.037
Observations	1119		1320		967		1470	
R squared	0.854		0.928		0.918		0.932	
F test for joint significance of the three variables	0.242		0.002		0		0.159	
	San Francisco		San Jose		Washington, D.C.			
	beta	s.e	*beta*	*s.e*	*beta*	*s.e*		
Close to transit station	−0.143	0.026	0.136	0.048	0.045	0.022		
Close to walk and ride station	−0.002	0.060	−0.091	0.062	0.039	0.028		
Close to transit station*tract income above msa 1970 median	0.081	0.042	−0.282	0.058	0.009	0.026		
Observations	1650		1155		2525			
R squared	0.9520		0.9590		0.9190			
F test for joint significance of the three variables	0.000		0.000		0.007			

Note: The unit of analysis is a census tract. This table reports results from 15 separate regressions of Equation (1). In each regression, there are fixed effects for the tract, the calendar year and an interaction between the calendar year and the tract's distance from the CBD. For the pooled regression, this last variable is interacted with *Metro Area Fixed Effects*. The omitted category is a tract that is more than one mile from a rail transit station. *Close To Transit* is a dummy variable that equals one if a tract is within one mile of a rail transit station. *Close to Walk and Ride Station* is a dummy that equals one if a tract is within one mile of a rail transit station that does not have a parking lot. *Tract Income Above MSA 1970 Median* is a dummy variable that equals one if a tract's 1970 household income is greater than the median household income in the same metropolitan area in 1970. The sample includes all census tracts within 20 miles of the CBD of the metro areas listed in Table 2.

Jose it has a large negative effect. In Boston, the average tract treated with a "Park and Ride" station experienced a 5% decline in home prices, while the average tract treated with a "Walk and Ride" station experienced a 7% increase in home prices.[22] Table 5 highlights that some cities, such as Atlanta, Baltimore, Boston, Chicago, Dallas, Sacramento and Washington, DC, have experienced greater home price appreciation in new "Walk and Ride" communities than in new "Park and Ride" communities. Boston is the city exhibiting the strongest evidence of poverty magnet effects in new "Park and Ride" communities and gentrification in new "Walk and Ride" communities.

Recent real estate research has emphasized the importance of supply constraints in considering real estate market dynamics (Glaeser, Gyourko and Saks 2005). Put simply, if new rail transit station construction increases local demand for housing among the rich, but if zoning restrictions impede new housing supply then we will observe a greater increase in nearby home prices and less population growth than if housing supply were more elastically supplied.[23] Research using San Diego data has emphasized that, due to fiscal revenue considerations, areas near rail transit stops were mostly zoned commercial (Boarnet and Crane 2001).

While it is extremely difficult to construct comparable zoning information across my 14-city sample by census year, anecdotes suggest that zoning does affect equilibrium outcomes near new stations. In Washington, DC, condo developers have been blocked at suburban metro stops: "Concerned about dense new development close to his own Northern Virginia home, U.S. Rep. Tom Davis, R-11th, promised hundreds of angry constituents gathered at Oakton High School last month that he would use his congressional clout to block Metro from selling a three-acre parcel near the Vienna station to Pulte Homes, a Michigan-based developer."[24] This editorial continues on about the conflict between the incumbent suburbanites, developers and urban planners: "The Fairfax Board agreed to let Pulte (a developer) raze 61 older single-family homes and plop down 2,250 luxury condo and town house units (with prices ranging

---

[22] New "Park and Ride" stations may lead to gentrification in communities that are more than one mile from the closest new station but that are still a short driving distance to a new station. Such communities would gain greater access to public transit but would be far enough from the station to not be subject to localized congestion nuisances around the station. I have tested this hypothesis and find no evidence to support this claim.

[23] I have examined population growth near new rail transit stations. Across the 14-city sample, the average "Park and Ride" station community experiences a 9% reduction in population when it is treated relative to population growth in the same metropolitan area in control tracts that are equidistant to the CBD. In contrast, the average new "Walk and Ride" station experiences 4% extra population growth relative to control tracts.

[24] Examiner Editorial - Fairfax Board: 'Let them eat condos' http://www.dcexaminer.com/articles/2005/05/10/opinion/editorial/85edit09condos.txt.

between $400,000 and $800,000, clearly not 'affordable housing') in 12-story towers right next to Interstate 66 without a full traffic study, just a vague idea that thousands more people can be stuffed onto Orange Line trains that are already jam-packed during rush hour. Supervisors also agreed to sell thousands of existing Metro parking spaces to make room for the new development's 4,400 additional cars. This is not 'smart growth.' This is madness."[25]

Additional Gentrification Results

A key indicator of gentrification is the share of college graduates living in the community. In Table 6, I report estimates of Equation (1) based on this outcome measure. Across the whole 14-city sample, a "Park and Ride" treatment reduces the share of adults who are college graduates by 1.9 percentage points. A "Walk and Ride" treatment in a tract with above-MSA median income and above-MSA median population density increases the share of college graduates by 5.1 percentage points ($-0.019 + 0.034 + 0.036$).[26] Across the 14 cities, Boston stands out as a city whose estimate of the "Walk and Ride" effect (B_2 in Equation (1)) is large. Running the same regression using the log of household income as the dependent variable yields similar qualitative effects, but the magnitudes of the coefficients are smaller. Across the whole 14-city sample, a "Park and Ride" treatment lowers tract household income by 2%, while a "Walk and Ride" treatment raises tract income by 4%. This difference between the two treatment types is statistically significant at the 1% level.[27]

Stratifying across the 14 cities, there has been a positive and statistically significant increase in the share of adults who are college graduates in nine of the cities in communities treated with a "Walk and Ride" station access. In contrast, in eight of the cities, communities treated with a "Park and Ride" station experienced a statistically significant reduction in the share of adults who are college graduates. The results in Table 6 support the hypothesis that the introduction of "Walk and Ride" stations induces gentrification, while "Park and Ride" stations attract the poor to live nearby.

---

[25] Examiner Editorial - Fairfax Board: 'Let them eat condos' http://www.dcexaminer.com/articles/2005/05/10/opinion/editorial/85edit09condos.txt.

[26] In results that are available on request, I have followed Brueckner and Rosenthal (2006) and included measures of the age of the census tract's housing stock in the regressions. Using data from the year 2000, my estimates support their point that more educated people are less likely to cluster in communities with an older housing stock. Controlling for the housing stock age variables has little effect on changing the estimates of B_{m1} and B_{m2} from Equation (1).

[27] These results are robust to controlling for a dummy variable that equals one if the census tract's borders changed over time.

Table 6 ■ Share college graduate dynamics in treated census tracts.

The dependent variable is the share of census tract adults who are college graduates.	All 14 MSAs		Atlanta		Baltimore		Boston	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	−0.019	0.003	−0.033	0.015	−0.020	0.008	−0.002	0.011
Close to walk and ride station	0.034	0.004	0.091	0.020	0.100	0.013	0.140	0.017
Close to transit station*tract income above msa 1970 median	0.036	0.004	0.057	0.020	0.022	0.013	0.019	0.025
Observations	30195		1702		2069		1656	
R squared	0.924		0.903		0.927		0.946	
	Chicago		Dallas		Denver		Los Angeles	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	−0.033	0.010	−0.031	0.018	0.134	0.020	−0.077	0.008
Close to walk and ride station	0.070	0.013	0.099	0.020	−0.047	0.030	0.034	0.009
Close to transit station*tract income above msa 1970 median	0.054	0.014	−0.022	0.021	0.103	0.084	0.072	0.010
Observations	2663		2181		1961		6270	
R squared	0.942		0.906		0.909		0.914	

(Continued)

Table 6 ■ Continued

	Miami		Portland		Sacramento		San Diego	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Close to transit station	−0.051	0.009	−0.018	0.015	0.030	0.012	−0.054	0.008
Close to walk and ride station	−0.008	0.017	0.020	0.014	0.083	0.017	0.040	0.011
Close to transit station*tract income above msa 1970 median	0.145	0.016	−0.008	0.013	−0.087	0.017	0.037	0.012
Observations	1293		1489		1101		1697	
R squared	0.91		0.906		0.89		0.937	
	San Francisco		San Jose		Washington, D.C.			
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*		
Close to transit station	−0.034	0.009	0.040	0.014	0.020	0.008		
Close to walk and ride station	−0.005	0.017	−0.039	0.018	0.044	0.010		
Close to transit station*tract income above msa 1970 median	0.027	0.016	−0.083	0.019	0.045	0.010		
Observations	1889		1312		2912			
R squared	0.9390		0.9290		0.9200			

Note: The unit of analysis is a census tract. This table reports results from 15 separate regressions of Equation (1). In each regression, there are fixed effects for the tract, the calendar year and an interaction between the calendar year and the tract's distance from the CBD. For the pooled regression, this last variable is interacted with Metro Area Fixed Effects. The omitted category is a tract that is more than one mile from a rail transit station. *Close To Transit* is a dummy variable that equals one if a tract is within one mile of a rail transit station. *Close to Walk and Ride Station* is a dummy that equals one if a tract is within one mile of a rail transit station that does not have a parking lot. *Tract Income Above MSA 1970 Median* is a dummy variable that equals one if a tract's 1970 household income is greater than the median household income in the same metropolitan area in 1970. The sample includes all census tracts within 20 miles of the CBD of the metro areas listed in Table 2. With the exception of Portland, all of the coefficients listed in this table are jointly statistically significant at the 1% level. Portland's p value for the three explanatory variables' joint significance is 0.202.

Testing for Longer Adjustment Periods

The panel estimates of Equation (1) focus on tract dynamics over a 10-year window. This may be too short of a time interval for analyzing community dynamics. For example, consider the census tracts located within one mile of Boston's Davis Square station. The Davis Square Red Line station was built in the mid-1980s. The share of adults in this community in the year 1970 was 8.2%. This increased to 18.4% in 1980, 35.7% in 1990 and 49.7% in 2000. The panel estimator would focus on the change in college graduates before and after treatment (*i.e.*, in 1980 and 1990) but note the large increase in gentrification from 1990 to 2000. One plausible explanation for this ongoing rise is the social multiplier effect. Gentrification feeds on itself as schools improve, local amenities improve and shopping and cuisine options improve.

In Table 7, I report estimates of Equation (2).[28] In this regression, the outcome measure is defined as the census tract's change in levels between 1970 and 2000.[29] This regression is estimated for tracts that were more than a mile from rail transit in 1970.

$$\begin{aligned} Y_{jm2000} - Y_{jm1970} = \Phi_{mt} &+ B_{m1}(\mathit{Years\ Close\ to\ Transit}_{j2000}) \\ &+ B_{m2}(\mathit{Years\ Close\ to\ Walk\ \&\ Ride}_{j2000}) \\ &+ B_{m3}(\mathit{Years\ Close}_{j2000} * X_{jm1970}) + \mathit{controls} + U_{jmt}. \quad (2) \end{aligned}$$

Unlike in Equation (1), the key explanatory variables are not dummy variables. Instead, in Equation (2), the key explanatory variables represent the number of years that a tract has been exposed to transit. For example, a tract near the Davis Square stop in Boston would be assigned 15 years experience near a "Walk and Ride" stop in the year 2000.[30]

28 The danger with this approach is that, if the treatment tracts were experiencing gentrification trends before they were treated with transit, then this long difference will attribute all of the 1970 to 2000 change in home prices and increases in the share of college graduates to the treatment.

29 Knaap, Ding and Hopkins (2001) provide an additional justification for studying long differences. They provide evidence from Washington County in Oregon that plans for light rail investments have positive effects on land values in proposed station areas. If this is the case, then home prices will increase before the tract is treated. Panel estimates of Equation (1) would underestimate the price appreciation impact in this case because the expectation of treatment raises tract home prices before the tract is indeed treated. The long difference estimates of Equation (2) would be more likely to measure the total effect.

30 For tracts that experienced proximity to rail transit in the 1970s, I assigned them 25 years of access experience in the year 2000. Similarly, I assigned treated tracts in the

Table 7 ■ Adjustment lags and gentrification in treated communities.

	log(*Home Price*)		Percent College Graduates	
	beta	*s.e*	*beta*	*s.e*
Years tract within one mile of rail station	−0.002	0.001	0.001	0.000
Years tract within one mile of walk and ride station	0.007	0.002	0.003	0.001
Years close to transit station*tract income above msa 1970 median	0.002	0.002	−0.001	0.001
log(Home price in 1970)	−0.198	0.014		
Percent college graudates in 1970			0.334	0.014
Observations	6164		7205	
Adjusted R squared	0.374		0.175	
Distance to CBD	Yes		Yes	
Metropolitan area fixed effects	Yes		Yes	

Note: This table reports long differenced regressions. In the left column, the dependent variable is the log(*Home Price in 2000/Home Price in 1970*). In the right column, the dependent variable is the *Share of Adults who are College graduates in 2000 - Share of Adults who are College Graduates in 1970*. The data set includes tracts within 20 miles of the CBD.

Table 7 reports two estimates of Equation (2). Unlike the results reported in Tables 5 and 6, the regressions are not stratified by metropolitan area. The left columns report the home price regressions. All else equal, a tract that has been near a "Park and Ride" station for 10 years experiences a 1.9% reduction in home prices, while a tract that has been near a "Walk and Ride" station for 10 years experiences a 5.4% increase in home prices, while a tract that has been close to a "Walk and Ride" station for 20 years would experience a 10.8% increase in home prices (between 1970 and 2000), relative to control tracts.[31] The right column of Table 7 reports estimates of Equation (2) using the community's share of adults who are college graduates as the dependent variable. All else equal, a tract that has been near a "Park and Ride" station for 10 years experiences a 1.1% point increase in the share of residents who are college graduates, while a tract that has been near a "Walk and Ride" station for 10 years experiences a 4.2% point increase in the share of residents who are college graduates.

1980s 15 years of experience and tracts treated in the 1990s with 5 years of transit access experience in the year 2000.

[31] The control tracts are census tracts in the same metropolitan area that have never been within a mile of rail transit station and are the same distance to the CBD as the treated tract.

Contrasting OLS and Instrumental Variables Results

In closing this section, I contrast instrumental variables estimates with ordinary least squares (OLS) estimates. Throughout this article, I have argued that communities are more likely to be treated if they anticipate that their community will benefit from such treatment. This suggests that OLS estimates of B_{m1} and B_{m2} in Equation (1) should be larger than instrumental variables estimates of these coefficients because the latter ideally represent the effect for a randomly treated community. OLS estimates of treatment effects should be larger than instrumental variables estimates because the latter ideally represent random assignment of treatment.

To compare OLS and instrumental variables (IV) results, I estimate 14 separate regressions of Equation (3) for the subset of census tracts that were more than one mile from the closest rail transit station in 1970.

$$Y_{j2000} - Y_{j1970} = \Phi + B * Close\ to\ \ Rail\ Transit_{j2000} + controls + U \quad (3)$$

Note that in Equation (3), I do not disaggregate rail stations into "Walk and Rides" and "Park and Rides." I do not have two credible instruments for these different treatments. Thus, I aggregate them into one category. In the IV regressions, I instrument using a tract's distance to the closest railroad line and the census tract's log of population density in 1970 (recall the results reported in Table 4 demonstrating their statistical significance).[32]

In Table 8, I report 56 estimates of B. It is important to emphasize that this estimate represents a weighted average of the "Walk and Ride" effect and the "Park and Ride" effect. Recall that for Boston, the former had a positive gentrification impact while the latter had a negative effect. Based on the OLS results, only three cities (Boston, Chicago and Washington, DC) experienced gentrification in the treated communities based on changes in college graduate shares. In 10 of the 14 cities, the IV point estimate is smaller than the OLS estimate.[33] The income regression results in columns 3 and 4 reveal roughly similar patterns. An explanation for the difference in the OLS and IV results is

[32] The coefficient results I report below are generally robust to excluding tract population density in 1970 as an instrument but the R^2 in the first stage of two-stage least squares declines if I exclude this variable.

[33] Boston, Chicago, Los Angeles and San Jose stand out as the cities whose IV results are larger than the OLS results, and these estimates in the college graduate regression are statistically significant.

Table 8 ■ Contrasting ordinary least squares and instrumental variables estimates.

	Percent of Adults who are College Graduates				log(Average Household Income)			
Column	(1)		(2)		(3)		(4)	
	beta	*s.e*	*beta*	*s.e*	*beta*	*s.e*	*beta*	*s.e*
Atlanta	0.008	0.026	**−0.814**	**0.436**	0.051	0.054	−3.144	1.881
Baltimore	−0.008	0.017	**−0.251**	**0.079**	0.019	0.032	**−0.552**	**0.181**
Boston	**0.077**	**0.015**	**0.181**	**0.109**	0.035	0.030	**0.425**	**0.260**
Chicago	**0.039**	**0.010**	**0.409**	**0.172**	**0.071**	**0.025**	**0.736**	**0.336**
Dallas	−0.020	0.026	−0.667	0.668	0.003	0.058	−0.544	0.739
Denver	0.062	0.040	−0.866	0.478	0.121	0.078	−1.833	1.195
Los Angeles	**−0.038**	**0.008**	**0.229**	**0.095**	**−0.074**	**0.020**	0.015	0.192
Miami	**−0.038**	**0.019**	**−0.164**	**0.064**	−0.013	0.057	−0.044	0.174
Portland	**−0.045**	**0.014**	**−0.600**	**0.142**	**−0.117**	**0.028**	**−1.109**	**0.284**
Sacramento	−0.008	0.019	**−0.125**	**0.061**	−0.046	0.045	**−0.236**	**0.141**
San Diego	−0.018	0.014	−0.032	0.025	**−0.075**	**0.029**	0.041	0.060
San Francisco	−0.025	0.014	**−0.170**	**0.037**	**−0.067**	**0.029**	**−0.312**	**0.101**
San Jose	**−0.035**	**0.017**	**0.171**	**0.090**	**−0.055**	**0.029**	**0.285**	**0.148**
Washington, DC	**0.095**	**0.015**	**−0.146**	**0.061**	**0.071**	**0.027**	**−0.231**	**0.100**
Estimation Method	OLS		IV		OLS		IV	

Note: This table reports estimates of *B* from 56 separate regressions based on Equation (3) in the text. The dependent variable is the 1970 to 2000 change. The sample includes all tracts that were more than one mile from a rail transit station in 1970 and that are within 20 miles of the metro area's CBD. The instrumental variables are a polynomial of a tract's year 2000 distance to the closest railroad line and the census tract's log of population density in 1970. The omitted category is the set of census tracts that were more than one mile from the nearest rail station in 1970 and 2000. In each regression, I control for the tract's distance to the CBD and the tract's value for the dependent variable in 1970. Statistically significant results are reported in bold.

that there are heterogeneous returns to rail transit access.[34] Those communities that gain the most from such treatment are more likely to be treated (Heckman, Urzua and Vytlacil 2006).

Conclusion

Over the last 30 years, several major cities have made huge investments in constructing and upgrading their rail transit systems. The investments described in Table 1 have created a large number of communities that are now within one

[34] The instrumental variables approach predicts tract treatment status based on the tract's distance to a preexisting railroad line (a cost shifter).

mile of new rail stations. Access to such stations, especially the "Walk and Ride" stations, increases the supply of transit-oriented communities where people can live, commute and shop while using the private auto less frequently. This article has used a census tract panel data set to examine demographics dynamics in such treated communities relative to observationally identical control communities. I have not attempted to measure how entire metropolitan areas are affected by new rail transit construction.[35]

Across the 14-city sample, new transit's local impacts differ significantly. Some cities, such as Boston and Washington, DC, have experienced gentrification in communities with increased access to rail transit, especially communities treated with a new "Walk and Ride" station. In other cities, such as Los Angeles and Portland, I find no evidence of gentrification based on home price dynamics and shares of the community that are college graduates. Unlike "Walk and Ride" stations, communities close to new "Park and Ride" stations often experience increases in poverty.

In considering the real estate impacts brought about by local public goods improvements, it is important to emphasize that I have only examined residential demographics dynamics. Future work might examine changes in commercial real estate activity near new transit stations. For example, do high-quality commercial stores open near new "Walk and Ride" transit stations? Do employers who employ workers who commute using public transit cluster near new stations (see Holzer, Quigley and Raphael 2003)?

I thank two reviewers and seminar participants at MIT, USC and Wharton, and I thank Nate Baum-Snow, Dora Costa, Ed Glaeser, David Luberoff and Bill Wheaton for helpful comments.

References

Altshuler, A. and D. Luberoff. 2003. *Mega Projects: The Changing Politics of Urban Public Investment*. Washington, DC: Brookings Institution Press.

[35] Banzhaf *et al*. (2004) present an applied general equilibrium analysis of the sorting pattern induced by ambient air quality shifts in the Los Angeles area. In their empirical application, the Clean Air Act shifts the Los Angeles distribution of air pollution. Optimizing agents respond to these changes in ambient air quality by relocating to their new utility-maximizing community. Housing prices across Los Angeles communities adjust such that, in equilibrium, aggregate demand for housing in each community equals the aggregate supply. They show that home owners in communities that always had clean air, such as those close to the Pacific Ocean, could be made *worse off* due to the effectiveness of the Clean Air Act because clean air is a less scarce commodity in the post-regulation period. In contrast, in my partial equilibrium study, I am implicitly assuming that communities that are more than one mile away from rail transit stations are not affected by the treatment. In my 14-city sample, only a small fraction of census tracts in each metropolitan area have been treated (see Table 1).

Banzhaf, S., H. Sieg, V.K. Smith and R. Walsh. 2004. Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods. *International Economic Review* 45: 1047–1077.

Baum-Snow, N. and M. Kahn. 2005. The Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970 to 2000. G. Burtless and J.R. Pack, editors. *Brookings-Wharton Papers on Urban Affairs* 2005: 147–206.

Bayer, P., F. Ferreira and R. McMillan. 2005. Tiebout Sorting, School Multipliers and the Demand for School Quality. NBER Working Paper #10871.

Boarnet, M. and R. Crane. 2001. *Travel by Design*. New York: Oxford University Press.

Bond, E. and E. Coulson. 1989. Externalities, Filtering, and Neighborhood Change. *Journal of Urban Economics* 26: 231–249.

Bowes, D. and K. Ihlanfeldt. 2001. Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics* 50: 1–25.

Brueckner, J. and S. Rosenthal. 2006. Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns be Rich? Working Paper. University of California-Irvine.

Cervero, R. and J. Landis. 1997. Twenty Years of the Bay Area Rapid Transit System: Land Use and Development Impacts. *Transportation Research Part A* 31: 309–333.

Coulson, N. and M. Lahr. 2005. Gracing the Land of Elvis and Beale Street: Historic Designation and Property Values in Memphis. *Real Estate Economics* 33: 487–507.

Gibbons, S. and S. Machin. 2005. Valuing Rail Access Using Transport Innovations. *Journal of Urban Economics* 57: 148–169.

Glaeser, E. and M. Kahn. 2004. Sprawl and Urban Growth. V. Henderson and J. Thisse, editors. *Handbook of Urban Economics*, Vol. 4. Amsterdam: North Holland Press.

Glaeser, E., M. Kahn and J. Rappaport. 2007. Why Do the Poor Live in Cities? The Role of Public Transportation. *Journal of Urban Economics*. Forthcoming.

Glaeser, E., J. Gyourko and R. Saks. 2005. Why is Manhattan So Expensive? Regulation and the Rise in House Prices. *Journal of Law and Economics* 48: 331–370.

Heckman, J. and E. Vytlacil. 1999. Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. *Proceedings of the National Academy of Sciences* 96: 4730–4734.

Heckman, J., S. Urzua and E. Vytlacil. 2006. Understanding Instrumental Variables in Models with Essential Heterogeneity. *The Review of Economics and Statistics* 88: 389–432.

Holzer, H., J. Quigley and S. Raphael. 2003. Public Transit and the Spatial Distribution of Minority Employment: Evidence from a Natural Experiment. *Journal of Policy Analysis and Management* 22: 415–442.

Knaap, G., C. Ding and L. Hopkins. 2001. Do Plans Matter? The Effects of Light Rail Plans on Land Values in Station Areas. *Journal of Planning Education and Research* 21: 32–39.

McMillen, D. and J. McDonald. 2004. Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983–1999. *Real Estate Economics* 32: 463–486.

Moffit, R. 1999. Models of Treatment Effects When Responses are Heterogeneous. *Proceedings of the National Academy of Sciences* 96: 6575–6576.

Tu, C. and M. Eppli. 1999. Valuing New Urbanism: The Case of Kentlands. *Real Estate Economics* 27: 425–451.