2004 V32 3: pp. 463–486
REAL ESTATE
ECONOMICS

Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983–1999

Daniel P. McMillen* and John McDonald**

This study examines the effect of the new rapid transit line from downtown Chicago to Midway Airport on single-family house prices before and after the opening of the line. The results show that the housing market anticipated the opening of the line. House prices were being affected by proximity to the stations in the late 1980s and early 1990s—after the plans for the line were well known. The difference between the increase in the value of homes within the sample area as compared with properties farther away from the new transit stations was approximately $216 million between 1986 and 1999.

The Midway Rapid Transit Line opened in Chicago on October 31, 1993. Also known as the Orange Line or Southwest Side Rapid Transit Line, it connects downtown Chicago to Midway Airport, a distance of 11 miles. The purpose of this paper is to measure the effects of the new line on the market for single-family houses. The data for the study are 17,034 single-family house transactions within 1.5 miles of the line that took place over the 1983–1999 period. This data set is far more comprehensive than data sets used previously to study the effects of transportation improvements. Use of data over the 17-year period permits us to estimate the effect of the line on house prices as it varies over time. Our study is also unique in its use of both the hedonic approach and the repeat sales estimator to track temporal variations in transit station gradients, and in using the repeat sales method to obtain estimates of the aggregate benefit of the new transit line.[1]

*University of Illinois at Chicago, Department of Economics and Center for Urban Real Estate, Chicago, IL 60607 or mcmillen@uic.edu.
**University of Illinois at Chicago, Department of Economics and Center for Urban Real Estate, Chicago, IL 60607 or mcdonald@uic.edu.

[1] Several studies have been conducted to estimate the impact of a rapid transit line on house prices; these studies include Mudge (1972), Boyce *et al*. (1972), Allen and Mudge (1974), Dewees (1976), Damm *et al*. (1980), Bajic (1983), Voith (1993), Gatzlaff and Smith (1993), McDonald and Osuji (1995) and Baum-Snow and Kahn (2000). However, except for Gatzlaff and Smith, these studies do not use both repeat-sales and hedonic methods, and only Voith estimates the effect of rapid transit lines over time. McDonald and Osuji estimate the effect of the Midway line on residential land values in 1990

We find that the anticipated benefits of the new transit line began to be capitalized into house prices as early as 1987, 6 years before construction was completed. The house price gradient with respect to distance from the nearest transit station rose in absolute value from 4.2% before 1987 to 19.4% during 1991–1996. However, we also find a decline in the gradient to 9.8% during 1997–1999 as home prices rose more rapidly in locations farther from the transit stations. On average, homes in our sample area increase in value by about $6,000 as compared with similar properties at the sample boundary, 1.5 miles from the new transit line. We estimate the aggregate increase in property values to be $215.9 million in 1997 dollars, which is nearly half the cost of constructing the transit line.

The Midway Line

The Midway Line provides the first rapid transit service to the southwestern quadrant of the city of Chicago. The other sectors of the city have all had rapid transit service since the turn of the 20th century, but the southwest side had been omitted. Although discussions of a possible southwest side elevated line had taken place since the 1940s, the transportation plan produced in 1962 by the Chicago Area Transportation Study (CATS, 1959, 1961, 1962) once again omitted the southwest side line. Instead CATS proposed express bus service for the Stevenson expressway, the highway that serves the southwest side, and an elevated line that would have run due east from Midway Airport to the old line that runs from downtown to the south side of the city. See McDonald (1988) for a retrospective examination of the original CATS plan.

New discussions of a southwest side line were initiated in 1977, and the intention to build a line was announced by Mayor Jane Byrne of Chicago in 1979 at the same time she announced the cancellation of an expressway project known as the "Crosstown Expressway." At the time, it was presumed that federal funds that had been earmarked for this expressway project would be transferred to the transit line, but this transfer of funds did not begin for another 5 years. As Krueger *et al.* (1980) and Anas (1982, 1983) discuss in detail, several alternative alignments for the proposed line were under consideration during the 1980–1983 period. The effects of various proposed alignments on the housing market were simulated by Anas (1982, 1983); this study is discussed below. A decision on the alignment was made, and on June 26, 1984, the first $32 million for property acquisition was received from the federal government.

compared with 1980, and they find that the land market had anticipated the opening of the line.

Table 1 ■ Midway line average weekday ridership in 2000 by station.

Station	No. of Riders per Day
Midway Airport	8,531
Pulaski	4,924
Kedzie	2,965
Western	3,268
35th/Archer	2,218
Ashland	1,253
Halsted	2,337
Roosevelt	2,609
Total	28,015

The Midway Line was constructed at a cost of $410 million, and the expense for rolling stock was an additional $111 million. The Federal Transit Administration provided 85% of the cost, and the State of Illinois provided the remaining 15%. The line is 11 miles in length, of which 9.2 miles are newly built track (including 2.7 miles of elevated structure). The line has eight stations outside the downtown area, and the fare is $1.50 for a one-way trip. Parking is available at the five stations located the greatest distances from downtown. The demand for these spaces immediately exceeded the supply at the original price of $1.50 per day. Typical travel time from Midway Airport to downtown is 30 minutes. Express bus service for the same trip during peak periods requires 45–50 minutes. Ridership on the line (*i.e.*, passengers entering the system at one of the Midway Line stations) was projected to be about 25,000 riders per weekday. Actual ridership in the first year was 28,000 per weekday, and average weekday ridership in 2000 by station is listed in Table 1. Note that 48% of the riders entered at one of the two stations that are most distant from downtown, and 30% of the passengers entered at Midway Airport.

Theoretical Considerations

The Midway line is a new transit line that replaces bus service to and from Midway Airport and other locations and also draws commuters who used the private auto. Does a transportation improvement of this magnitude require the construction of a general equilibrium model of transportation and the housing market, or will a partial equilibrium model be sufficient? Some earlier studies, such as those by Boyce *et al.* (1972) and Allen and Mudge (1974), have concluded that increases in property values near a new transit line are accompanied by noticeable property value decreases in other locations. The interpretation of the empirical results in this study is based on the presumption that only a partial equilibrium model is needed. The Midway Line creates improved access to downtown Chicago and other destinations around eight station sites. The housing market in the vicinity of the stations adjusts, and it shall be assumed that the general equilibrium effects on the rest of the metropolitan area can be ignored. The general equilibrium effects on the housing market are likely to be small and diffuse within a metropolitan area of over seven million people.

The validity of the assumption that a partial equilibrium model is sufficient in this case was, in effect, tested by Anas (1982, 1983). Anas develops a large-scale general equilibrium model of transportation, housing and residential location for metropolitan Chicago that is used to simulate the effects of the Midway Line on the housing market (and on modal choice as well). One of the options for the Midway Line that is tested is very similar to the one that was actually built. The simulations for this option (known as the IHB project) show negligible effects on housing rent beyond a distance of 1.5 miles from the stations.[2]

Another theoretical question concerns the effect of the transportation improvement on house prices prior to the opening of the transit line. Assume that the transit project is announced at some point in time, and it is completed j years later. In the simplest case the change in value P due to the transit line of a particular house at a time prior to opening is

$$\Delta P_{t-k} = \frac{\Delta R_t(d)}{r(1+r)^k}, \tag{1}$$

where t is the year in which the transit line opens, $t-k$ is the year under study ($k \leq j$ years prior to t), r is the real discount rate and ΔR_t is the change in rent from time t forward, which is a function of distance d to a transit station. The facts about the transit line—station locations, travel time and fares—are known with certainty beginning at time $t-k$, and the date of completion is expected to be t. Equation (1) simply says that the increase in house price is discounted back to time k. Housing rent does not rise before opening day, but house price does rise in anticipation of the increase in housing rent, which is caused by the improvement in accessibility. Note that the elasticity of ΔP_{t-k} with respect to $E\Delta R_t$ is unity, and that P increases from time $t-j$ up to time t at rate r.

In this simple case, prices begin to rise at the time the project is announced, after which they rise at rate r up to the time the project is completed. Real prices are then constant thereafter. In the more realistic case of uncertainty, expectations may eventually prove to be incorrect. After the project is completed, prices may continue to adjust if the value of the project was assessed inaccurately beforehand. For example, the benefit from being close to the station may be

[2] See Anas (1982) for a full description of the model and McDonald and Osuji (1995) for a summary of the simulation results.

overestimated before the line is constructed, leading to a drop in the gradient after the line has been in place from some time. Ultimately, the time path of the station gradient is an empirical issue.

Data Sources

Data on the sales prices of all houses sold in Cook County, Illinois, for the years 1983–1999 (with the exception of 1992) were provided by the Illinois Department of Revenue from transaction tax records. Houses were identified by the Property Identification Number (PIN) so that the price could then be matched with the characteristics of the house as recorded for 1997 by the Cook County Assessor. The street address of each house was located in a geographic information system (GIS) program as follows. The PIN identifies the quarter section in which the property that was sold is located (*i.e.*, an area one-half mile square). The Cook County Assessor provides the street address of the property *owner*, and this address can be coded into the GIS program. The street address of the property owner is assumed to be the same as the street address of the property itself if the address is located within the quarter section noted by the Illinois Department of Revenue. For all of Cook County, geographic coordinates for approximately 80% of the houses were identified using this method. However, Chicago has a lower success rate using this method than the rest of Cook County because it has more rental properties; 47.8% of the homes in the quarter sections represented in the data set had to be dropped from the analysis.[3] The location of each house was then used to measure distance to the nearest Midway Line station, the nearest point on the Midway Line and downtown Chicago. Census tract data from the 1990 Census of Population and Housing were also added to each observation using the GIS program.

House transactions were selected for the study using the following criteria:

- The house is located within 1.5 miles of the Midway Line.
- The house is located closer to a Midway Line station than to a station for another transit line.
- The house is located within 1 of 12 community areas that are listed in Table 2.
- The house is not located west of Midway Airport.

[3] Values do not differ substantially for the two sets of properties. The average assessed value in 1997 was $9,228 for the 17,034 homes in the final sample, compared with $9,289 for the remaining 15,591 houses. Although single-family homes in Chicago are supposed to be assessed at 16% of property value, the average ratio of price to assessed value was 10.4 for the homes that sold in 1997. Thus, these average assessed values translate into market values of $95,971 and $96,606.

Table 2 ■ Descriptive statistics.

Variable	Full Sample	Repeat Sales
Price	77462.191	89602.485
	(32324.162)	(29469.961)
	[1400, 262501]	[5825, 222375]
Lot size	3669.017	3667.831
(square feet)	(785.280)	(740.565)
	[378, 16758]	[702, 9392]
Building area	1059.584	1051.515
(square feet)	(273.245)	(263.894)
	[400, 3952]	[400, 2976]
Age	59.318	62.555
	(25.056)	(23.892)
	[0, 131]	[0, 130]
Number of bedrooms	2.613	2.582
	(0.689)	(0.678)
	[1, 7]	[1, 6]
More than one story	0.178	0.201
Masonry used in construction	0.599	0.620
Basement	0.683	0.701
Finished basement	0.181	0.178
Attic	0.461	0.490
Finished attic	0.111	0.114
Central air conditioning	0.117	0.113
Garage, 1 car	0.308	0.314
Garage, 2 or more cars	0.448	0.454
Proportion Hispanic in census tract	0.238	0.240
	(0.168)	(0.161)
	[0.000, 0.940]	[0.000, 0.940]
Proportion Black in census tract	0.057	0.036
	(0.194)	(0.143)
	[0.000, 1.000]	[0.000, 1.000]
Proportion vacant housing in census tract	0.049	0.047
	(0.038)	(0.034)
	[0.000, 0.229]	[0.000, 0.229]
Archer Heights	0.058	0.054
Bridgeport	0.039	0.035
Brighton Park	0.109	0.111
Chicago Lawn	0.033	0.041
Clearing	0.052	0.058
Gage Park	0.217	0.245
Garfield Ridge	0.088	0.079
McKinley Park	0.040	0.032
New City	0.077	0.055
West Elsden	0.123	0.126
West Englewood	0.009	0.003
West Lawn	0.157	0.161

Table 2 ■ continued.

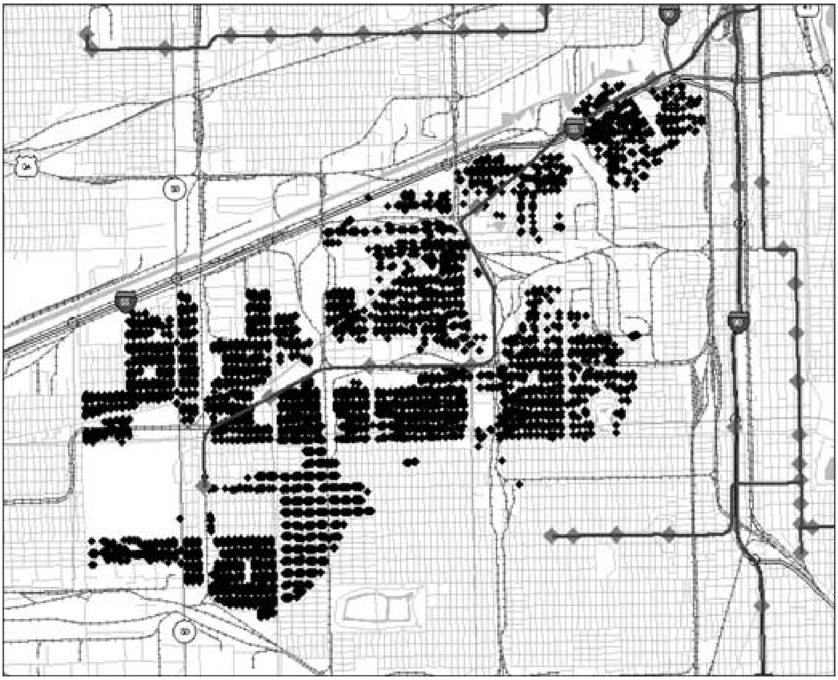
Variable	Full Sample	Repeat Sales
1983	0.037	0.000
1984	0.040	0.001
1985	0.045	0.006
1986	0.055	0.014
1987	0.061	0.020
1988	0.064	0.035
1989	0.069	0.054
1990	0.065	0.067
1991	0.060	0.075
1992—No observations	0.000	0.000
1993	0.068	0.081
1994	0.077	0.097
1995	0.076	0.106
1996	0.067	0.094
1997	0.070	0.104
1998	0.069	0.108
1999	0.078	0.140
Within 1/8 mile of transit line	0.044	0.042
Distance from CBD	7.189	7.268
	(1.530)	(1.486)
	[2.572, 10.054]	[2.577, 10.054]
Distance from transit station	0.811	0.811
	(0.331)	(0.331)
	[0.034, 1.737]	[0.034, 1.732]

Note: The sample average is presented in all cells. For continuous variables, the standard deviation is presented in parentheses, followed by the minimum and maximum in brackets. The full sample has 17,034 observations, and the repeat-sales sample has 4,056 observations.

This selection method resulted in a sample of 17,034 individual transactions and 4,056 repeat sales observations within the area so defined. The location pattern of the transactions data is shown in Figure 1, which also shows the Midway Line and its stations. The clusters of homes around the line show clearly that our sample is drawn from an area around the full transit line, rather than around the stations.

The characteristics of the sample are shown in Table 2. The mean selling price for the full sample is $77,462, while the mean selling price in the repeat-sales sample is $89,602. The mean house in the sample is 59 years old and has 1,060 square feet, a lot size of 3,669 square feet and 2.6 bedrooms. Masonry construction was used in 60% of the houses, 68% have a basement and 75%

Figure 1 ■ Sample area.



have a garage.[4] Each house is located in 1 of 12 community areas listed in Table 2. Table 2 also shows the proportion of the observations from each year for 1983–1999; unfortunately, no data are available for 1992. Also, note that the proportion of the repeat-sales sample increases over time because of the nature of the method used to construct a repeat-sales sample. A larger proportion of observations drawn from later years leads to higher mean values for sales prices and ages in the repeat sales sample. Finally, Table 2 shows that the mean distance from a Midway Line station is 0.81 miles, the mean distance from downtown Chicago is 7.2 miles, and 4.4% of the observations are within 1/8 mile (one city block) of the Midway Line.

[4] One weakness of our data set is that all housing characteristics are recorded in a single year, 1997. The assumption that housing characteristics are constant over time forms part of the basis for the repeat-sales estimator, but unfortunately we have no means of checking for violations of the assumption. Some of the apparent price appreciation may be due to changing home quality if homeowners remodel their houses as prices rise. Changes in home quality may also lead to a bias in our estimated gradients if the probability of remodeling falls with distance from the transit stations. Unfortunately, we cannot avoid this problem, which is endemic to both the repeat-sales and hedonic estimators.

Econometric Methods

Our objective is to estimate the time path of the effect of distance from the nearest transit station on house prices. Since the station sites were known well in advance, significant gradients may appear long before the new transit line opened. Our estimating equations must be sufficiently flexible to account for this temporal variation in the transit gradients.

The base estimating equation is the standard hedonic price function:

$$\ln P_{it} = \alpha_t + \delta_t d_i + \beta' X_i + u_{it}, \tag{2}$$

where P_{it} is the sales price of house i, which sold at time t, d_i is the distance of house i from a transit station, X_i is a vector of housing characteristics and u_{it} is an error term. The estimated values of α_t form the house price index. These values are the estimated coefficients for a set of dummy variables indicating the quarter of sale, with $\alpha = 0$ for the base quarter, 1983:1. Similarly, the estimated transit station gradient index is the set of coefficients for interaction terms between the quarterly dummies and distance from the station that is nearest to the house once the new transit line opened.

The estimated transit station index proves to be highly volatile when the index is allowed to vary freely across years. It may be more reasonable to assume that the gradients are constant for several successive years, with break points between some years. Following the switching regression approach used by Brueckner (1986) and McMillen (1994), we use the Akaike (1973) and Schwarz (1978) information criteria to choose the optimal number of break points.[5] We allow the coefficients for three spatial variables—distance to the nearest transit station, a dummy variable indicating that a house is within 1/8 mile (a standard Chicago block) of the transit line and distance from the city center—to vary across regimes; other coefficients are constrained to be constant over time. Both the AIC and SC criteria produce three identical break points of 1986/87, 1990/91 and 1996/97.

We also use the repeat-sales method to estimate the indexes. The repeat-sales method was proposed by Bailey, Muth and Nourse (1963) as a method for constructing house price indexes. It can be derived directly from Equation (2). For those houses that sell more than once during a sample period the change in $\ln P$ can be written

[5] The AIC test (Akaike 1973) chooses the number of terms in the expansion by minimizing $AIC(k) = \ln \hat{\sigma}^2 + 2k/n$, where k is the number of estimated coefficients, $\hat{\sigma}^2$ is the estimated variance of the error term and n is the sample size. The Schwarz criterion (Schwarz 1978) chooses k to minimize $SC(k) = \ln \hat{\sigma}^2 + 2 \ln k/n$. In general, the optimal number of terms will be larger under the Akaike criterion. However, in our application, both criteria lead to the same choice of k.

$$\ln P_{it} - \ln P_{is} = \alpha_t - \alpha_s + d_i(\delta_t - \delta_s) + u_{it} - u_{is}, \tag{3}$$

where s is the time of the previous sale. Equation (3) is a repeat-sales estimator that expresses the percentage change in the price of a house as a function of an implicit set of time-specific explanatory variables. The series of coefficients $\alpha_1, \ldots, \alpha_T$ forms the price index. The time-specific coefficients (δs) of distance to the rapid transit station show how the value of this feature varies over time. Note that additional variables with time-varying coefficients can be added to Equation (3) in the same way that the coefficients for d_i are estimated: Each additional variable is interacted with the standard series of discrete variables used to estimate the αs.

By controlling for the effect of omitted variables that do not change over time, the repeat-sales estimates are potentially subject to less bias than standard hedonic estimates. However, the reduction in bias comes at the expense of a large decrease in sample size and possible selection bias.[6] As noted above, the reduction in sample size in this study is from 17,034 to 4,056. The repeat-sales estimator has important advantages for our study. Our key variable of interest, distance from the new transit stations, is correlated with other important spatial variables, such as distance from the city center. Misspecification of the functional form for these variables—or their omission—will bias hedonic estimates of the δs. However, these variables do not change over time. If their coefficients are also constant over time, then the repeat-sales estimator provides unbiased estimates of the index of the δs. We test whether the coefficients for other spatial variables (distance from the city center and a location within 1/8 mile of the transit line) vary over time by including them in our repeat-sales estimator. F-tests do not reject the null hypothesis that only the coefficients for distance to the transit stations vary over time.

McMillen and Dombrow (2001) proposed an extension of the standard repeat-sales estimator that helps mitigate the effects of smaller sample sizes by imposing a smooth price index. Let T_{it} represent the month of sale for house i, and let T_{is} equal its previous month of sale. McMillen and Dombrow use a smooth continuous function $g(T)$ to represent the time trend in house prices, replacing the series of coefficients $\alpha_1, \ldots, \alpha_T$ in Equation (3). Similarly, the function $h(T)$ replaces the δ coefficients. The repeat-sales estimator becomes

$$\ln P_{it} - \ln P_{is} = g(T_{it}) - g(T_{is}) + d_i[h(T_{it}) - h(T_{is})] + u_{it} - u_{is}. \tag{4}$$

Following Gallant (1981, 1982), McMillen and Dombrow (2001) use a Fourier expansion to model $g(T_{it})$ and $g(T_{is})$. The Fourier expansion uses sine and cosine terms to approximate a function.

[6] See McMillen and Dombrow (2001) and McMillen (2003) for a discussion of these issues.

The first step is to transform the time variable to lie between 0 and 2π; define $z_{it} = 2\pi T_{it}/\max(T)$ and $z_{is} = 2\pi T_{is}/\max(T)$, where $\max(T) = 204$ months. We again use the AIC and Schwarz information to guide our choice of expansion length for the Fourier expansion. Both information criteria lead to the following expansions:

$$\begin{aligned} g(T_{it}) - g(T_{is}) = {} & \gamma_1(z_{it} - z_{is}) + \gamma_2\left(z_{it}^2 - z_{is}^2\right) + \gamma_3(\sin(z_{it}) - \sin(z_{is})) \\ & + \gamma_4(\cos(z_{it}) - \cos(z_{is})) + \gamma_5(\sin(2z_{it}) - \sin(2z_{is})) \\ & + \gamma_6(\cos(2z_{it}) - \cos(2z_{is})) \end{aligned} \tag{5}$$

$$\begin{aligned} h(T_{it}) - h(T_{is}) = {} & \lambda_1(z_{it} - z_{is}) + \lambda_2\left(z_{it}^2 - z_{is}^2\right) + \lambda_3(\sin(z_{it}) - \sin(z_{is})) \\ & + \lambda_4(\cos(z_{it}) - \cos(z_{is})) \end{aligned} \tag{6}$$

These estimated coefficients can then be used to form estimates of the price index and the transit station gradient index.[7]

Empirical Results

The repeat-sales data are used to estimate both Equation (3), where the discrete time periods are quarters, and the Fourier expansion model. The results are depicted in Figures 2 and 3. Figure 2 presents the quarterly nominal repeat-sales price index for 1983:1 to 1999:4 as the jagged line. The Fourier nominal repeat-sales price index is the middle smooth curve. The upper and lower smooth curves represent 95% confidence intervals for the Fourier price index. Note that the quarterly index and the Fourier index track closely together, although the quarterly index has appreciable jumps that are unlikely to represent real market behavior accurately. House prices increased by 90% over the 1983–1999 period (85% up to 1998). These results are comparable to those presented by McMillen (2003) for the entire city of Chicago, which show a 97% increase from 1983 to 1998. McMillen also shows that the price increase on the southwest side was somewhat below the overall city increase.

The most important results in the study are depicted in Figure 3, which shows the quarterly and Fourier indexes for the price gradient with respect to distance to the transit station. The quarterly index moves erratically, but it clearly shows a negative gradient after the opening of the Midway line in late 1993 compared with the initial value in 1983 on the order of −10% to −15% per mile. The

[7] See McMillen and Dombrow (2001) and McMillen (2003) for details.

Figure 2 ■ Repeat sales and Fourier price indexes.

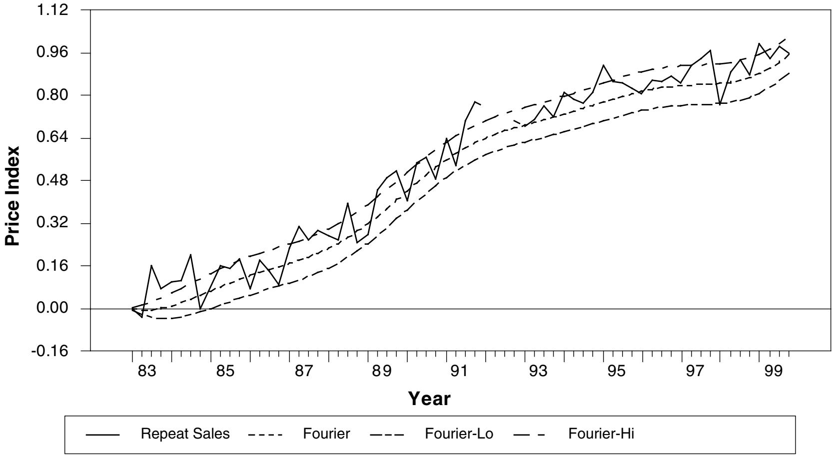
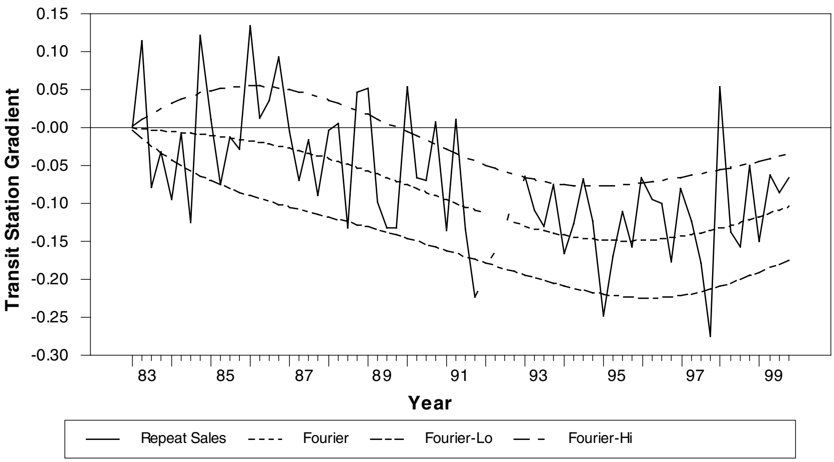


Figure 3 ■ Repeat-sales and Fourier confidence intervals for transit station gradient index.



Fourier index shows a smooth change in this price gradient to −15% per mile in 1995 (compared with the initial value in 1983) and then a modest reduction in the negative gradient to −11% per mile in 1999. Figure 3 also shows that the 95% confidence interval for the Fourier index falls below zero (compared with the initial value in 1983) in late 1989, which is 4 years prior to the opening of the Midway Line. A change in the gradient of −15% per mile means, for example, that a house located at 1.25 miles from a station sold for 15% less than did a house located 0.25 miles from a station. Given a mean house price of $89,000 in the repeat-sales sample, this price difference is $13,350.

Figures 2 and 3 clearly show the advantages of the Fourier approach. The Fourier estimates produce realistic smooth temporal changes in house prices and transit stations gradients. In contrast, the standard repeat-sales estimates produce erratic indexes with sharp changes between quarters. Neither prices nor gradients actually jump sharply between March 31 and April 1, as implied by the standard estimator. The smooth Fourier estimates provide a more realistic representation of changes in prices and gradients over time, and smooth indexes such as those in Figures 2 and 3 make it easier to track historical trends in the housing market.

The repeat-sales method clearly shows that house prices reacted to the Midway Line by generating a significant price gradient with respect to distance to the stations. The hedonic method is also used as a check on this conclusion and because the results are of interest. As noted above, Gatzlaff and Smith (1993) use both the repeat-sales and the hedonic methods in their study of the impact of the Miami Metrorail on house prices, but they find only weak effects. The hedonic model (Equation (2) above) includes the housing and neighborhood characteristics listed in Table 2, dummy variables for the 12 community areas, dummy variables for each year and variables to capture the effects of

- distance to downtown Chicago,
- location within one block of the Midway Line and
- distance to the transit station.

It turns out that, according to F-tests, the effects of these three variables vary significantly over time. The regression results are shown in Table 3.

Consider first the coefficients of the standard hedonic variables. The empirical results are that the characteristics of the house and the neighborhood nearly all have the expected effects. These variables include the lot size, the building area, age of the structure, number of bedrooms, whether the house has more than one story (a negative effect), masonry construction, basement and finished basement (omitted category is no basement), attic and finished attic (omitted category is no attic), central air conditioning (an unexpected zero impact), garage, proportion of population in census tract that is Hispanic or Black (strongly negative effects) and the housing vacancy rate in the census tract (a negative effect). The coefficients of the dummy variables for community areas (where the omitted community area is Albany Park) show that house prices vary appreciably across neighborhoods for reasons that are not accounted for

Table 3 ■ Hedonic regression results and specification tests.

Variable	Coefficient	T-value	Variable	Coefficient	T-value
Constant	7.044	52.249	*Clearing*	−0.232	−11.403
ln lot size	0.219	17.973	*Gage Park*	−0.053	−4.200
ln building area	0.251	21.053	*Garfield Ridge*	−0.064	−4.467
Age	−0.005	−38.165	*McKinley Park*	0.211	8.560
Number of bedrooms	0.021	5.220	*New City*	−0.094	−4.328
More than one story	−0.078	−11.465	*West Elsden*	−0.090	−7.962
Masonry	0.021	3.324	*West Englewood*	−0.213	−6.930
Basement	−0.010	−1.570	*West Lawn*	−0.083	−5.518
Finished basement	0.028	4.923	*Near El Line, base*	−0.127	−4.794
Attic	0.001	0.177	*Near El Line* × *87–90*	0.109	3.187
Finished attic	0.042	5.107	*Near El Line* × *91–96*	0.128	3.981
Central air conditioning	−0.009	−1.250	*Near El Line* × *97–99*	0.094	2.749
Garage, 1 car	0.057	9.825	*Distance to CBD, base*	0.069	8.820
Garage, 2 or more cars	0.074	12.671	*Distance to CBD* × *87–90*	0.017	3.554
Proportion Hispanic	−0.533	−19.141	*Distance to CBD* × 91–96	0.0003	0.068
Proportion Black	−0.475	−22.084	*Distance to CBD* × *97–99*	−0.022	−4.559
Proportion Vacant	−0.578	−5.400	*Distance to transit station* × *base*	−0.042	−2.245
Bridgeport	0.403	12.779	*Distance to transit station* × *87–90*	−0.074	−3.292
Brighton Park	0.119	7.765	*Distance to transit station* × *91–96*	−0.152	−7.052
Chicago Lawn	−0.049	−2.764	*Distance to transit station* × *97–99*	−0.056	−2.363
$R^2 = 0.699$					

Number of observations = 17,034.
The dependent variable is the natural logarithm of the sales price.
Note: The regression also includes 15 dummy variables indicating the year of sale.

by the other variables included. Not surprisingly, Bridgeport is the community area with the highest house prices, *ceteris paribus*. Bridgeport was the home of the first Mayor Daley, was the home of the current Mayor Daley until a few years ago and is generally regarded as a neighborhood with powerful local political connections.

Next, consider the results pertaining to distance to downtown Chicago (CBD). Separate price gradients have been estimated for four time periods: 1983–1986, 1987–1990, 1991–1996 (with 1992 omitted) and 1997–1999. These four time periods were determined endogenously using statistical tests that are described in the previous section. The results in Table 3 show that the price gradient with respect to distance to the CBD was +6.9% per mile in the 1983–1986 time period. The coefficients for the other time periods refer to the change in the price gradient relative to this initial value. In particular, the price gradient for 1987–1990 increased by 1.7% relative to the initial value, the gradient for 1991–1996 returned to the initial value, and then the gradient declined by 2.2% relative to the initial value in 1997–1999. The empirical results for the effect of being within one block of the Midway Line are presented in similar fashion. The effect of such a location in the initial period of 1983–1986 was −12.7%. In the three later periods this coefficient is very close to zero because the coefficients report changes relative to the initial value. For example, in 1987–1990 the coefficient changes by +10.9% to yield a net value of −1.8%.

The most important findings in Table 3 are those for distance to the transit station. The results show that there was a negative price gradient of −4.2% in the initial period of 1983–1986. The price gradient became more negative by −7.4% in the 1987–1990 period, was even more negative by −15.2% in 1991–1996 and then flattened to −5.6% relative to the initial gradient in 1997–1999. These results confirm the time pattern observed in Figure 3 for the change in the price gradient with respect to the transit stations. As a summary, Figure 4 shows the quarterly repeat-sales, Fourier and hedonic model estimates of the change in this price gradient. All three indexes move together rather closely, except that the hedonic index for the 1997–1999 period is above the Fourier index by about 6%. Figure 3 again shows the advantages of the Fourier approach in tracking historical movements in the gradients. The smooth Fourier index provides a much more realistic depiction of the temporal gradient movements than the erratic hedonic estimates.

Table 4 reports the F-tests that were used to test whether the effects of being within one block of the Midway line, distance to the CBD and distance to the transit station vary over time. The table shows that all three effects vary over time in the hedonic model, but that only the effect of distance to the transit station varies over time in the quarterly repeat-sales model. The results

Figure 4 ■ Repeat-sales, Fourier and hedonic transit station gradient index.

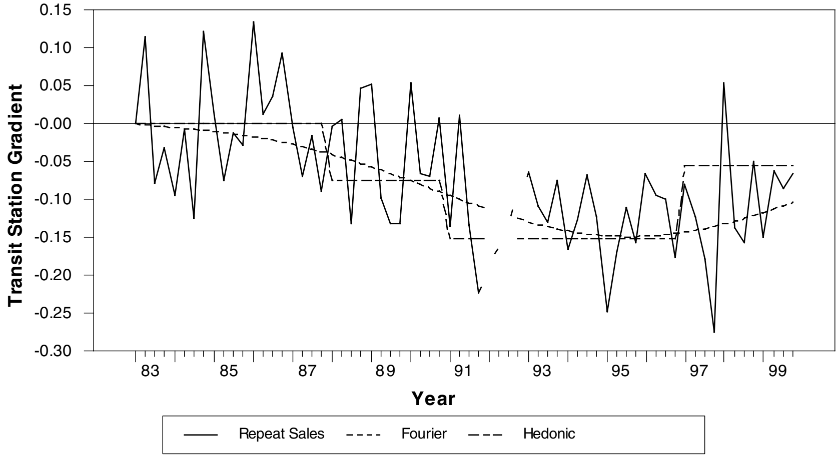


Table 4 ■ *F*-tests for constant coefficients over time.

Coefficients	*F*-Test Value	Degrees of Freedom	*P*-Value
Hedonic			
Proximity to transit line	5.519	3, 16979	0.001
Distance to CBD	25.357	3, 16979	0.000
Distance to transit station	18.573	3, 16979	0.000
Repeat sales			
Proximity to transit line	0.970	62, 3805	0.544
Distance to CBD	1.125	63, 3805	0.234
Distance to transit station	1.770	63, 3805	0.000
Proximity to transit line and distance to CBD	1.004	125, 3805	0.471

Note: For the repeat-sales estimator, the degrees of freedom are lower for proximity to the transit line because there were no sales of houses within 1/8 mile of the line in one quarter.

that have been reported in Figures 2, 3 and 4 are based on the findings in Table 4.

Our results so far are based on the assumption that the effect of distance from the nearest transit station does not vary across the eight stops. However, we have reason to suspect that this assumption is overly restrictive. The first two stations after the CBD have no parking. The last stop on the line is Midway Airport, and commuters must compete with airline passengers for parking spaces. Thus, a logical grouping of stations is (1) the Midway stop, (2) the middle four stations

Table 5 ■ Estimated transit station gradients.

	1983–86	1987–90	1991–96	1998–99
Distance to transit station				
Base model, pooled observations	−0.042	−0.117	−0.194	−0.098
	(2.245)	(7.262)	(13.277)	(5.606)
Midway station	−0.002	−0.097	−0.158	−0.042
	(0.069)	(3.889)	(7.163)	(1.665)
Middle four stations	0.052	−0.125	−0.236	−0.107
	(2.890)	(7.972)	(16.754)	(6.391)
Two stations nearest CBD	−0.010	0.024	0.017	0.019
	(0.932)	(2.270)	(1.730)	(1.825)
Proximity to transit line				
Base model, pooled observations	−0.127	−0.019	0.001	−0.033
	(4.794)	(0.859)	(0.032)	(1.509)
Gradients vary by station	−0.084	−0.020	−0.010	−0.034
	(3.195)	(0.909)	(0.571)	(1.575)

Note: Absolute T-values are in parentheses. The base model results are derived from Table 3. The remaining results are derived from a regression that mimics the base model, with separate variables for the three sets of transit stations.

and (3) the two stations nearest the CBD, which do not have parking. We use the hedonic model, which is presented in Table 3, as our base specification, while allowing for separate gradients across the four time periods and across the three station types. Due to a small number of observations within a block of the tracks, we continue to impose the restriction that the discount for proximity to the transit line does not vary over time.

Pertinent results from the extended model are presented in Table 5. The first row of results reproduces estimates that can be derived from Table 3: The transit station gradient was 4.2% from 1983 to 1986, 11.7% from 1987 to 1990, 19.4% from 1991 to 1996 and 9.8% thereafter. The estimates from the extended model show that significant gradients are confined to the stations with parking. The gradients are largest for the four middle stations, where the bulk of the passengers are commuters. For the middle four stations, we find that prices actually *increased* with distance from the future transit stations sites from 1983 to 1986. Overall, the pattern of gradients over time is similar to the estimates from the base model for the stations with parking, and the estimated discount for proximity to transit line is about the same as before.

Figures 5–7 show the estimated relationships between the price index and distance from the three sets of transit stations at four different times. The lines are normalized such that the mean for the set of predicted values equals the actual

Figure 5 ■ Distance from the Midway station.

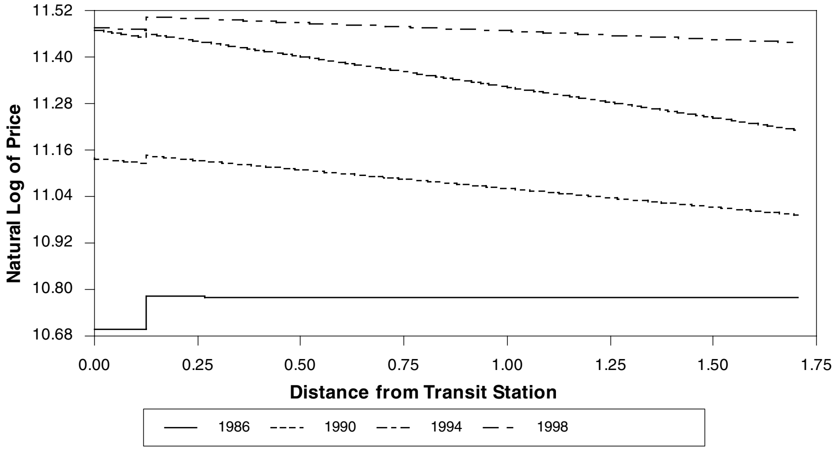
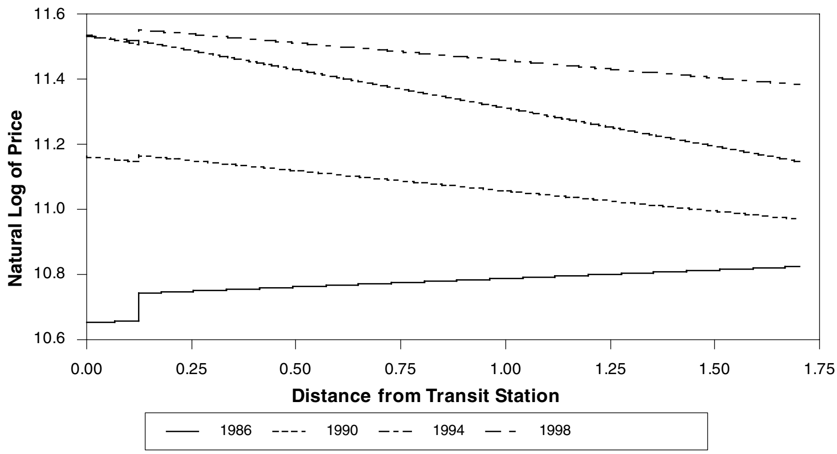
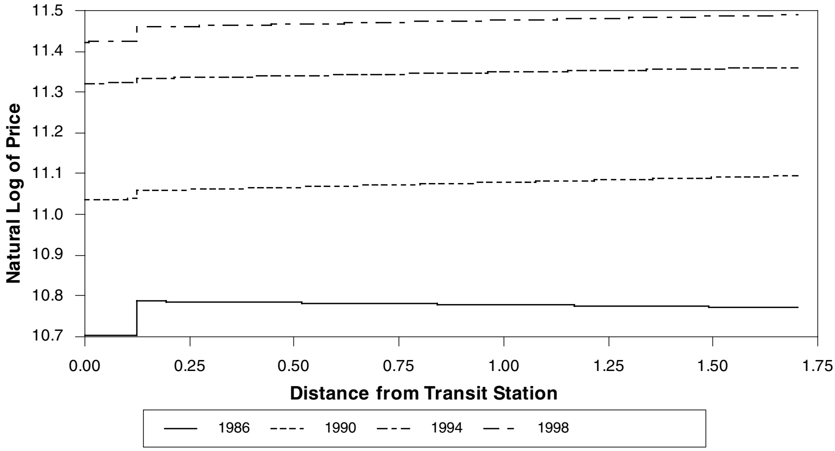


Figure 6 ■ Distance from the middle four stations.



mean value for the natural logarithm of prices in that year. Figures 5 and 6 show that prices rose much more rapidly during 1986–1990 and 1990–1994 near the Midway stop and near the other stations with parking lots. Between 1994 and 1998, prices increased more rapidly farther away from these transit stations. The only important difference between Figures 5 and 6 is the upward slope for the 1986 function for the middle four transit stations. In contrast, prices near the transit stops closest to the CBD were largely unaffected by the new transit line: The price indexes shown in Figure 7 are nearly parallel to each other.

Figure 7 ■ Distance from the two stations closest to the CBD.



Figures 5–7 clearly show the significance of the estimated discount for being within a block of the transit line during the initial time period. In later times, the discount nearly disappears. This result appears anomalous: why should prices be lower near a nonexistent and unanticipated transit line? In fact, the result is not surprising because the new transit line was routed through existing rail rights of way and through industrial areas. The relatively few homes near the future transit line were thus in unattractive locations. Figure 1 shows that these homes also tended to be near transit stations. Much of the discount for sites near the tracks was eliminated once the new line became anticipated as formerly undesirable locations became attractive due to their ready access to the new stations. It appears that any negative externalities of locations close to the line such as noise and traffic congestions are overwhelmed by the advantages of access to the transit system, although the relatively small number of homes near the transit line prevents this claim from being conclusive.

Figures 5 and 6 show clearly that from 1986 to 1990 and again to 1994, the sharpest rises in prices came close to the stations. Changing expectations likely account for the contrary finding that prices rose more rapidly at more distant locations between 1994 and 1998. Initial expectations appear to have been that most of the benefit of the new line would accrue to homes close to the stations. Since parking is not common along Chicago's rapid transit system, it is possible that prices increased more rapidly in more distant locations after 1993 because the benefits of the new parking lots became clear. However, this explanation is speculative; what is clear is that the sharp increase in prices near the stations came during the time before the transit line actually opened.

Estimating the Aggregate Increase in Property Values

The hedonic model implies that the new transit line was anticipated as early as 1987, with discrete changes in the gradient at the beginning of 1991 and 1997. From Equation (3), the estimated rate of appreciation between time t and some previous time s is simply $\alpha_t - \alpha_s + d_i(\delta_t - \delta_s)$. This expression would provide a direct estimate of the rate of appreciation induced by the new transit line if the new stations were the only factor leading to price increases. However, all of our figures are in nominal terms. The BLS price index for metropolitan Chicago is not suitable for our purposes because during this time house prices rose much more rapidly than the general rate of inflation in the City of Chicago and the rate of appreciation varied widely across neighborhoods (McMillen 2003). To estimate how much prices rose because of the new transit line, we need a base appreciation rate for comparable locations that are not affected by the new line.

The edge of our sample—sites 1.5 miles from the transit line—is a suitable base location. While the new transit line can be expected to have a relatively small effect on property values at the edge of the sample, these locations are otherwise similar to the rest of the sample. The estimated appreciation rate for sites at the edge of the sample is given by Equation (3), with d_i evaluated at 1.5. The difference in appreciation rates between sites at $d_i = 1.5$ and sites closer to the new transit stations is simply

$$(d_i - 1.5)(\delta_t - \delta_s). \tag{7}$$

Our estimates of $\delta_t - \delta_s$, which are shown in Figure 3, are negative after late 1989. Since $d_i \leq 0$ in our sample, Equation (7) implies that appreciation rates rose more rapidly within our sample area than at the edge. Equation (7) may underestimate the effect of the new transit line on property values if properties 1.5 miles from the new stations also increased in value. However, the bias should be small since our results imply that property values decline markedly with distance from rapid transit stations.[8]

We can use any of our three estimation procedures—repeat-sales, Fourier repeat-sales or the hedonic approach—to construct an estimate of the aggregate

[8] The motivation for our approach comes from Thibodeau (1990), who estimates the effects of proximity to a high-rise office building on residential property values in Dallas. He compares his hedonic price estimates to predictions for the same property if it were 3000 meters away from the high rise. His specification is also similar to ours in that he allows for a discount for homes that are very close (within 100 meters) to the high rise. However, our approach differs in that we allow the transit station gradient to vary smoothly over time, which makes the repeat-sales approach a more natural basis for our estimates.

Table 6 ■ Aggregate increase in home values.

Time Interval	Average Difference in Appreciation Rates (%)	Aggregate Change in Homes Values (Millions of 1997$)
1986–90	4.02	125.9
1990–94	4.55	142.7
1994–99	−1.68	−52.7
1986–99	6.89	215.9

increase in property values induced by the new transit line. The Fourier estimates are clearly preferable to the standard repeat-sales approach because smoothed values are not as sensitive to the dates chosen for t and s. The Fourier estimates are also preferable to the hedonic estimates because they are not as prone to missing variable bias and do not impose the unrealistic assumption that the station gradient is constant over long time periods. Thus, our estimate of the aggregate increase in property values is derived from the Fourier model.

We base our comparisons on four different dates, which are derived from the switching regression model: 1986, the last year before distance from the new stations affected home values; 1990, the end of the interval during which the future stations first affected home values; 1994, the first year after the line was completed and 1999, the end of our sample period.

The estimates are based on the first quarter of each of these years.[9] To estimate the aggregate increase in property values, we must use assessed values since most homes do not sell in a given year. All of our assessed values are from 1997. We translate the assessed values into market prices using the sample average factor of 10.4. Finally, to account for the 15,591 homes that were not geocoded, we factor our estimate up again by $1.92 = (15{,}591 + 17{,}034)/17{,}034$. Although we established earlier that the average assessed values of the two sets of homes are not substantially different, we now must make an additional assumption that the distribution does not vary by distance from the nearest transit station. After these adjustments, the aggregate market value of all homes in the sample area is approximately $3.1 billion (1997$).

The results are shown in Table 6. For 1986–1990, the average value of Equation (7) across all properties in the sample is 4.02, which implies that

[9] The necessary figures for the calculations are $\delta_{90} - \delta_{86} = -0.05826$, $\delta_{94} - \delta_{90} = -0.06604$ and $\delta_{99} - \delta_{94} = 0.02437$.

the rate of increase of homes in the sample area was 4.02 percentage points higher than comparable homes 1.5 miles from the nearest transit station. Based on 1997 values, this 4.02% difference in growth rates translates into a difference in market values of $125.9 million. The differential increase in appreciation rates continues over the next time period, which includes the time when the transit line began operation. However, there is some backtracking between 1994 and 1999 as homes at the edge of the sample area appreciate more rapidly than homes in the interior. All told, Table 6 implies that appreciation rates were 6.89 percentage points higher between 1986 and 1999 than comparable homes 1.5 miles from the nearest transit station, translating into an aggregate increase in property values of $215.9 million. Using the BLS Chicago price index, this estimate is approximately $193.8 million in 1993 dollars. Thus, this price increase is approximately 47% of the construction cost of $410 million.

Summary and Conclusions

This study has employed a large sample of single-family house sales as well as both the repeat-sales method and the hedonic method to estimate the impact of the Midway Line on house prices. Both methods produce empirical results showing that the opening of the transit line in late 1993 had been anticipated in the housing market beginning in the late 1980s and perhaps as early as 1987. This finding confirms earlier results by Damm *et al*. (1980) for the Washington Metro and those by McDonald and Osuji (1995) for residential land values near the Midway Line in 1990. The study also shows that the house price gradient with respect to distance to the Midway Line stations became strongly negative after the line opened. The hedonic results show that the house price gradient with respect to distance to the station sites had been −4.2% per mile before 1987 and was −19.4% during 1991–1996, a change of −15.2%. The repeat-sales method yields an identical result for this change in the price gradient. Finally, both the repeat-sales results and the hedonic results show that this price gradient moderated somewhat after 1996. Perhaps there had been some overshooting as the housing market adjusted to the new transit line.

Despite this overshooting of home values, we find that home prices increased substantially due to the new transit line. Using 1986 as our base, we find that appreciation rates were approximately 6.89 percentage points higher between 1986 and 1999 than comparable homes farther from the nearest transit station. This difference in appreciation rates translates into an aggregate increase in property values of $215.9 million, or an average of about $6,000 per home. Our results thus provide strong evidence that the new transit line was highly valued by neighboring homeowners.

References

Akaike, H. 1973. Information Theory and the Extension of the Maximum Likelihood Principle. B. Petrov and F. Csaki, editors. *Second International Symposium on Information Theory*. Akailseoniai-Kudo: Budapest.

Allen, B. and R. Mudge. 1974. The Impact of Rapid Transit on Urban Development: The Case of the Philadephia Lindenwood High Speed Rail Line. RAND Corporation paper P-5246: Santa Monica, CA.

Anas, A. 1982. *Residential Location Markets and Urban Transportation*. Academic Press: New York.

———. 1983. The Effects of Transportation on the Tax Base and Development of Cities. U.S. Department of Transportation Report Number DOT/OST/P-30/85/005: Washington, DC.

Bailey, M., R. Muth and H. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association* 58: 933–942.

Bajic, V. 1983. The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto. *Urban Studies* 20: 147–158.

Baum-Snow, N. and M. Kahn. 2000. The Effects of New Public Projects to Expand Urban Rail Transit. *Journal of Public Economics* 77: 241–263.

Boyce, D., B. Allen, R. Mudge, P. Slater and A. Isserman. 1972. Impact of Rapid Transit on Suburban Residential Property Values and Land Development. Working Paper, NTIS 220693. Department of Regional Science, University of Pennsylvania.

Brueckner, J. 1986. A Switching Regression Analysis of Urban Population Densities. *Journal of Urban Economics* 19: 174–189.

Chicago Area Transportation Study. 1959, 1961, 1962. Chicago Area Transportation Study: Final Report (3 volumes). CATS: Chicago.

Damm, D., S. Lerman, E. Lerner-Lamm and J. Young. 1980. Response of Urban Real Estate Values in Anticipation of the Washington Metro. *Journal of Transport Economics and Policy* 14: 315–336.

Dewees, D. 1976. The Effect of a Subway on Residential Property Values in Toronto. *Journal of Urban Economics* 3: 357–369.

Gallant, A. 1981. On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form. *Journal of Econometrics* 15: 211–245.

———. 1982. Unbiased Determination of Production Technologies. *Journal of Econometrics* 20: 285–323.

Gatzlaff, D. and M. Smith. 1993. The Impact of the Miami Metrorail on the Value of Residences near Station Locations. *Land Economics* 69: 54–66.

Krueger, C., C. Heramb, B. Kunze and M. Gallery. 1980. Southwest Transit Study, Phase I Report: Preliminary Alternatives Analysis. City of Chicago, Department of Public Works, Bureau of Transportation Planning and Programming.

McDonald, J. 1988. The First Chicago Area Transportation Study Projections and Plans for Metropolitan Chicago in Retrospect. *Planning Perspectives* 3: 245–268.

McDonald, J. and C. Osuji. 1995. The Effect of Anticipated Transportation Improvement on Residential Land Values. *Regional Science and Urban Economics* 25: 261–278.

McMillen, D. 1994. Vintage Growth and Population Density: An Empirical Investigation. *Journal of Urban Economics* 36: 333–352.

———. 2003. Neighborhood House Price Indexes in Chicago: A Fourier Repeat Sales Approach. *Journal of Economic Geography* 3: 57–73.

McMillen, D. and J. Dombrow. 2001. A Flexible Fourier Approach to Repeat Sales Price Indexes. *Real Estate Economics* 29: 207–225.

Mudge, R. 1972. The Impact of Transportation Savings on Suburban Residential Property Values. Ph.D. dissertation, University of Pennsylvania.

Schwarz, G. 1978. Estimating the Dimension of a Model. *Annals of Statistics* 6: 461–464.

Thibodeau, T.G. 1990. Estimating the Effects of High-Rise Office Buildings on Residential Property Values. *Land Economics* 66: 402–408.

Voith, R. 1993. Changing Capitalization of CBD-Oriented Transportation Systems: Evidence from Philadelphia. *Journal of Urban Economics* 30: 360–372.