Available online at www.sciencedirect.com
ScienceDirect
Journal of Urban Economics 63 (2008) 816–840
JOURNAL OF Urban Economics
www.elsevier.com/locate/jue

Old homes, externalities, and poor neighborhoods. A model of urban decline and renewal

Stuart S. Rosenthal *

Department of Economics, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA

Received 10 March 2006; revised 4 June 2007

Available online 15 August 2007

Abstract

This paper investigates urban decline and renewal in the United States using three panels that follow neighborhoods on a geographically consistent basis over extended periods of time. Findings indicate that change in neighborhood economic status is common, averaging roughly 13 percent per decade; roughly two-thirds of neighborhoods studied in 1950 were of quite different economic status fifty years later. Panel unit root tests for 35 MSAs indicate that neighborhood economic status is a stationary process, consistent with long-running cycles of decline and renewal. In Philadelphia County, a complete cycle appears to last up to 100 years. Aging housing stocks and redevelopment contribute to these patterns, as do local externalities associated with social interactions. Lower-income neighborhoods appear to be especially sensitive to the presence of individuals that provide social capital. Many of the factors that drive change at the local level have large and policy relevant effects.

JEL classification: R11; R23; R31; N9

Keywords: Filtering; Externalities; Neighborhoods; Poverty

1. Introduction

Urban decline and renewal are events that evoke passionate responses from neighborhood residents and city officials alike. Yet although the economic vitality of neighborhoods is central to local public policy and urban development, evidence on the extent and nature of neighborhood decline and renewal is limited. This paper examines these issues. I begin with an observation.

Visit nearly any low-income urban neighborhood in the US and it is apparent that poor families occupy old homes built originally for higher income households.[1] This simple fact implies that neighborhoods cycle through regular periods of decline and renewal. These cycles arise because most existing housing stocks deteriorate over time until they are replaced with newer dwellings. Housing demand, meanwhile, increases with income. As a result, higher income families tend to move away from older neighborhoods while lower-income families move in. Contributing to this migration is a second quite different process based on local exter-

* Fax: +1 (315) 443 1081.
E-mail address: ssrosent@maxwell.syr.edu.
URL: http://faculty.maxwell.syr.edu/ROSENTHAL/.

[1] The primary exceptions are low-income families living in homes built through various federal and local government subsidy programs. With few exceptions, the private market does not build unsubsidized fixed site low-income housing (see Baer, 1986, for example).

0094-1190/$ – see front matter
doi:10.1016/j.jue.2007.06.003

Table 1
Transition probabilities of census tract relative income status between 1950 and 2000

	Low income in 1950	Lower-middle income in 1950	Upper-middle income in 1950	High income in 1950
(a) *Estimates are Based on 9419 Census Tracts from a Balanced Panel for 35 MSAs**				
Low income in 2000	34.21	17.90	17.86	16.36
Lower-middle income in 2000	28.31	26.42	26.04	18.17
Upper-middle income in 2000	23.84	32.12	26.94	21.50
High income in 2000	13.64	23.56	29.15	43.98
Total percent	100	100	100	100
(b) *Estimates are Based on 1191 Census Tracts from a Balanced Panel for PHILADELPHIA MSA**				
Low income in 2000	25.82	13.27	17.79	15.02
Lower-middle income in 2000	30.07	30.61	27.52	19.11
Upper-middle income in 2000	29.08	30.95	28.19	26.96
High income in 2000	15.03	25.17	26.51	38.91
Total percent	100	100	100	100
(c) *Estimates are Based on 1113 Census Tracts from a Balanced Panel for CHICAGO MSA**				
Low income in 2000	37.74	20.54	14.45	25.19
Lower-middle income in 2000	31.13	27.79	27.38	15.27
Upper-middle income in 2000	24.12	34.44	25.10	15.27
High income in 2000	7.00	17.22	33.08	44.27
Total percent	100	100	100	100

* Tracts with median income less than the city-wide 25th percentile in the given year are defined as low income. Tracts with median income between the 25th and 50th percentiles are defined as lower-middle income. Tracts with median income between the 50th and 75th percentiles are defined as upper-middle income. Tracts with median income above the 75th percentile are defined as upper income.

nalities. This occurs when individuals behave in ways that generate social capital and costs (as with gardening versus crime), and when status conscious households base their migration decisions on the attributes of their neighbors (as with race and education). Together, aging housing stocks and local externalities imply that change in neighborhood economic status should be common, at least over a sufficiently long time frame. Moreover, to the extent that change in neighborhood economic status is rooted in the nature of housing demand, aging of the housing stock, and social norms, it ought to be possible to anticipate and potentially to influence these events.

This paper explores these issues using a unique set of panels that follow individual neighborhoods on a geographically consistent basis over extended periods of time. A series of closely related questions are considered. Perhaps most fundamental, is change in neighborhood economic status common or rare? Do neighborhoods cycle up and down in economic status? Can change in neighborhood economic status be anticipated, and if so, what is the role of aging of the housing stock and redevelopment versus externalities arising from social interactions? What are the neighborhood factors that contribute most to change in neighborhood economic status? Each of these questions will be considered.

We begin by examining evidence on whether change in neighborhood economic status is common or rare. Table 1(a) through (c) sheds light on this issue. The table presents transition rates of neighborhoods between different levels of economic status using a balanced panel of census tracts for portions of 35 MSAs that are followed on a consistent geographic basis from 1950 to 2000.[2] Each census tract is treated as a separate neighborhood. Status is measured based on the average income of a neighborhood *relative* to the average income of all census tracts in the balanced panel for the MSA and year in which the neighborhood is observed. Neighborhood geography is coded to year 2000 census tract boundaries for all years. In addition, both here and elsewhere in the paper, neighborhoods are classified into four groups based on whether neighborhood relative income levels are in the first through fourth quartiles of relative income, referred to below as low income, lower middle-income, upper middle-income and high-income, respectively. Additional details on the data construction and sample composition are provided in the appendix.

A striking pattern emerges from all three portions of the table. For the 35 MSAs included in the sample, roughly two-thirds of all low-income neighborhoods in 1950 (Table 1(a), column 1) are of higher income status in 2000. In Philadelphia (Table 1(b)) and Chicago (Table 1(c)), similar magnitudes are evident. Among

[2] Only tracts assigned by Census to the same MSA in both 1950 and 2000 are included in the balanced panel. Details are in the appendix.

Table 2
Average absolute value of percentage change in census tract relative income by decade

Year	35 MSA Census Tract Balanced Panel 1950 to 2000	331 MSA Census Tract Panel 1970 to 2000
1950	–	–
1960	14.9	–
1970	13.3	–
1980	11.5	12.2
1990	13.9	13.1
2000	13.0	12.4

higher income tracts in 1950, change is also the dominant pattern in all three portions of the table. For all 35 MSAs (Table 1(a)), only 26.94 percent of upper-middle income tracts in 1950 were still of upper-middle income status in 2000. More generally, there is overwhelming evidence that most neighborhoods in the lower half of the income distribution tend to move up the economic ladder while most neighborhoods in the upper half tend to move down.

By how much does neighborhood economic status change with each passing decade? Table 2 provides evidence on this point by displaying the average absolute value of the change in Tract relative income on a decade-by-decade basis. Once again, a balanced panel of neighborhoods is used in constructing these measures. For the 35 cities followed from 1950 to 2000, the ten-year change in relative income status ranges from a high of 14.9 percent between 1950 and 1960, to a low of 11.5 percent between 1970 and 1980. From 1970 forward, all of the MSAs throughout the United States are followed using data obtained from Geolytics Inc. (see the Geolytics website, and the appendix, for details). For this expanded cross-section, the same pattern holds: in each decade, the average change in neighborhood relative income status is roughly 12 to 13 percent. These patterns make clear that change is the norm among urban neighborhoods in the United States. Yet, dramatic as these patterns are, to my knowledge this is the first time that these patterns have been reported. Possibly that is because most families remain in their homes (and neighborhoods) far less than ten years, a horizon too brief for the change in neighborhood economic status to be readily apparent to transient residents.[3]

Having established that most urban neighborhoods are in a constant state of transition, the remainder of this paper is organized as follows. Details on the neighborhood panels used for the analysis are provided in the appendix, along with additional summary measures. Although creation of the data was a major undertaking and there are many details, it is sufficient here to indicate that for each panel, all of the neighborhoods are followed on a geographically consistent basis over time, as in Tables 1 and 2. This leaves the body of the paper free to focus on analysis and results.

Section 2 explores the time series properties associated with change in neighborhood economic status. If neighborhoods cycle up and down around a stable long run mean, that would imply that neighborhood economic status is a stationary process as opposed to a random walk. This idea is considered by applying panel unit root tests to those census tracts followed from 1950 to 2000 on a decade by decade basis. Results strongly suggest that neighborhood economic status is a stationary variable. Additional analysis considers the degree of serial correlation in neighborhood economic status for period lengths ranging from 10 to 50 years. In all cases, results suggest that neighborhood cycles extend over many decades. For Philadelphia County, an area followed from 1900 to 2000, a typical neighborhood takes roughly 100 years to complete a full cycle of change up and down in relative income.

Section 3 elaborates on the mechanisms that drive neighborhood economic change including its cyclical nature. Two mechanisms are emphasized, aging of the housing stock along with periodic redevelopment—often referred to as the filtering process—and local externalities arising from social dynamics that are sensitive to the type of people who live in a community (e.g. homeowners, minorities, college educated). This section also describes the variables and develops the empirical model used to analyze the impact of these two mechanisms on neighborhood economic change.

Estimates from several variants of the empirical model are presented in Section 4 and are designed to shed light on two issues. First, to highlight differences in the manner that filtering and social dynamics affect neighborhood economic status. Second, to highlight the impact of individual control measures on neighborhood economic status, including the age distribution of the housing stock and different indicators of the community's socio-demographic composition (e.g. population age distribution, presence of minorities, homeownership, etc.). In all cases, the dependent variable in this portion of the paper is the change in neighborhood economic status between 1990 and 2000.

Results in Section 4 confirm that filtering and social dynamics affect neighborhood economic status through very different channels and with different temporal pat-

[3] The median renter, for example, moves roughly every one to two years, while the median homeowner moves in six to seven years (Rosenthal, 1988).

terns. Omitting socio-demographic (SES) controls from the regression model, for example, has little impact on the qualitative pattern of house age coefficients and only a modest effect on their magnitudes. This suggests that filtering and social dynamics have somewhat independent effects. Moreover, as the SES controls are more deeply lagged, their influence on a community's economic status attenuates sharply, but in contrast, the influence of the age distribution of the housing stock is quite persistent up to fifty years in the past. This pattern suggests that both SES factors and the age distribution of the housing stock affect change in neighborhood economic status, but in different ways.

To isolate the impact of the individual covariates I initially estimated a two-step GMM model treating all of the contemporaneous (year 1990) house age and SES controls as endogenous. For instruments I used house age and SES lags from 1980 and 1970, along with the lagged level of the neighborhood's economic status. Unfortunately, diagnostic tests fail to support the instrument strategy. In all of the models the instruments likely are weak and Hansen-J tests reject the overidentifying restrictions.[4] As an alternative I proceed as follows. Given my prior that 1970 and 1980 house age and SES attributes are exogenous to change in a neighborhood's economic status in the 1990s, rejection of the overidentifying restrictions implies that lagged neighborhood attributes belong in the structural model: in effect, history matters. Drawing on that idea, I focus on non-IV models using 1980 neighborhood attributes as covariates. Throughout, the dependent variable remains the change in neighborhood economic status in the 1990s.

Results are suggestive of a number of important relationships. With regard to filtering, controls are included for the age distribution of the housing stock in 1980 including housing 10 years or less in age, 10 to 19 years, on up to 30 to 39 years in age, and housing age 40 and over. Coefficients on these measures indicate that there is a U-shaped relationship between the age of the housing stock and change in a community's economic status in the 1990s. Moreover, the bottom of the "U" is centered on housing aged 10 to 19. This pattern is clearly present for neighborhoods in each of the upper three income quartiles in their MSAs. For the lowest income neighborhoods the pattern is more of a single step function with housing age 40 and over having the most positive impact on economic status in the 1990s. These patterns are consistent with the filtering/redevelopment model of housing markets. Specifically, on average, 1980 middle-aged housing decays with the passage of time but is not yet sufficiently old in the 1990s to warrant demolition and wholesale refurbishment. For that reason, middle-aged 1980 housing attracts lower income families in the coming decades. In contrast, the presence of older housing in 1980 is a predictor of urban redevelopment in the 1990s. Because higher income families are attracted to newly built housing, this contributes to an increase in a community's economic status in the 1990s and related gentrification.[5] These patterns also support the idea that filtering and redevelopment contribute to long-running cycles of neighborhood decline and renewal.

Among the socio-demographic indicators two dominant patterns emerge. First, various proxies indicate that the presence of individuals and conditions that foster the provision of local social capital impart positive spillovers on the community. Second, lower income neighborhoods appear to benefit most from such spillovers. This is especially evident with respect to the presence of homeowners, college educated individuals, and prime age workers, three groups that bring financial and human capital resources to a community. Results also indicate that high density development—a proxy for crime and other ills associated with congestion—hurts the economic status of a community in the 1990s, but most for the lower income communities. The presence of minorities (African American and Hispanic) has relatively little effect on the future status of lower-income communities, but is associated with a substantial decline in economic status among higher income neighborhoods in the 1990s. It is likely that this result reflects preferences for neighborhood racial composition, either because low-wealth minorities seek out locations already populated with such families, and/or because higher income white individuals flee neighborhoods where minorities are present.

The estimated effects also have economically important magnitudes. As an example, a 10 percentage point higher homeownership rate in 1980 would cause the economic status of a low-moderate income community to rise by 1.2 percentage points in the 1990s.[6] Similarly, a 10 percentage point higher presence of college ed-

[4] Details are provided later in the paper.

[5] Brueckner and Rosenthal (2006) develop a conceptual model that clarifies the spatial impact of aging of the housing stock on the location of low- and high-income neighborhoods and related patterns of gentrification. They also provide empirical evidence in support of the underlying conceptual framework.

[6] It is partly in the hope of such outcomes that both the Clinton and George W. Bush administrations have pushed to expand homeownership.

Table 3
Serial correlation in growth in neighborhood relative income (t-ratios based on robust standard errors in parentheses)

	Philadelphia County Ward Panel[a] $\log(y_{2000}/y_{1950})$	Philadelphia County Ward Panel[a] $\log(y_{1980}/y_{1950})$	Philadelphia County Census Tract Balanced Panel[b] $\log(y_t/y_{t-1})$	35 MSA Census Tract Balanced Panel[b] $\log(y_t/y_{t-1})$
$\log(y_{1950}/y_{1900})$	−0.9465			
	(−6.17)			
$\log(y_{1950}/y_{1920})$		−0.4538		
		(−3.20)		
$\log(y_{t-1}/y_{t-2})$			−0.0126	−0.0564
			(−0.24)	(−11.27)
Constant	−0.0617	−0.0423	−0.0863	–
	(−1.10)	(−0.82)	(−10.66)	
Period length (years)	50	30	10	10
Time span	1900 to 2000	1920 to 2000	1950 to 2000	1950 to 2000
County fixed effects	–	–	–	125
Observations	39	39	1304	37,676
R-square	0.4329	0.1749	0.0002	0.0323

[a] All data were coded to year 1900 Ward boundaries. See the appendix for details.
[b] All data were coded to year 2000 census tract boundaries. See the appendix for details.

ucated individuals in 1980 would cause the economic status of a low-income community to rise by 4.6 percentage points in the 1990s. A 10 percentage point increase in prime age individuals (age 30 to 55 in 1980) boosts the economic status of a low-income community by 5.1 percentage points relative to a corresponding increase in younger individuals. On the whole, these and other estimates are strongly suggestive that individuals with wealth and human capital provide social capital and in so doing generate positive spillovers that enhance the appeal of a neighborhood. A more complete description of these and other results is provided later in the paper. For now, we turn to the time series properties of neighborhood economic status.

2. Neighborhood economic cycles and mean reversion

It is clear that change in neighborhood economic status is the norm, not the exception. Accordingly, this section examines the time series properties of neighborhood economic status. As will become apparent, results shed further light on the cyclical nature of urban decline and renewal.

Consider first Table 3 which reports results from four regressions that characterize the degree to which change in neighborhood economic status is serially correlated. As before, neighborhood economic status (y_{it}) for neighborhood i in period t, is measured as the ratio of average income in the neighborhood relative to average income in its MSA.[7] The first three regressions focus on Philadelphia County. In the first two instances, the data are coded to the year 1900 Ward boundary geography, and there were 39 such Wards in Philadelphia in 1900. It should also be noted that income data for 1900 and 1920 is based on a constructed measure obtained from the IPUMs (see the appendix for details) and is subject to more measurement error than the income data from 1950 forward.[8] Bearing that in mind, in the first regression, the period length is set to 50 years, and $\log(y_{2000}/y_{1950})$ is regressed on $\log(y_{1950}/y_{1900})$. In the second regression the period length is set to 30 years, and $\log(y_{1980}/y_{1950})$ is regressed on $\log(y_{1950}/y_{1920})$. In the remaining two regressions in the table, the period length is set to 10 years, and $\log(y_t/y_{t-1})$ is regressed on $\log(y_{t-1}/y_{t-2})$. For these latter regressions, all data are analyzed at the census tract level coded to year-2000 census tract geographic boundaries, and tracts are followed from 1950 to 2000. The first of these regressions considers just census tracts in Philadelphia County while the second is

[7] To be precise, let y_{it} be the relative income of neighborhood i ($i = 1, \ldots, \mathrm{I}$) in period t. In addition, y_{it} is defined to be $Y_{it}/\bar{Y}_t$, where Y_{it} is the average level of income in tract i in period t, while $\bar{Y}_t$ is the city-wide average level of income in period t. By construction, the expected value of y over all neighborhoods in period t equals 1.

[8] Specifically, I used the IPUMs variable OCCSCORE to measure each individual's income in 1900 and 1920. That variable is an estimate of the income the individual would have earned in 1950 given their actual occupation.

based on census tracts included in the 35 MSA balanced panel.[9]

Moving from left to right in Table 3, the period length in the first three columns goes from longest to shortest, and always in reference to Philadelphia County. Observe that the coefficients on the lagged dependent variables are −0.95, −0.45, and −0.013, respectively, with t-ratios of 6.2 and 3.2 and 0.24. For the typical neighborhood in Philadelphia, a neighborhood's relative economic status in 2000 is 95 percent back to where it began 100 years earlier in 1900. As the period length narrows in the second and third regressions, neighborhood cycles are less complete and the coefficients on the lagged dependent variables are reduced, falling to 1.3 percent for the 10-year period length model. This latter estimate is close to the corresponding measure obtained for the 35 MSA panel in the last regression. In that model, the coefficient on the lagged dependent variable is −5.6 percent (with a t-ratio of 11.27). Together, these results lend support for the idea that neighborhood relative income cycles up and down over extended periods of time.

An implication of these results is that $\log(y_{it})$ has finite variance and a stable long-run mean even when t is large—in other words, $\log(y_{it})$ is stationary. This can be tested using panel unit root tests. To clarify, consider the following equation,

$$\log(y_{it}) = \theta_{i,o} + \theta_{i,1}\log(y_{i,t-1}) + e_{it}, \tag{2.1}$$

where $\log(y_{it})$ is expressed as a function of a constant and its one period lag, and i and t denote the individual census tract and time period as before. A necessary condition for $\log(y_{it})$ to be stationary is that $\theta_{i,1} < 1$. However, under the null that θ_1 equals 1, OLS estimates of (2.1) are biased.[10] Accordingly, subtracting $\log(y_{i,t-1})$ from both sides of (2.1),

$$\Delta\log(y_{it}) = \theta_{i,o} + \tilde{\theta}_{i,1}\log(y_{i,t-1}) + e_{it} \tag{2.2}$$

where $\Delta\log(y_{it}) \equiv \log(y_{it}) - \log(y_{i,t-1})$ and $\tilde{\theta}_{i,1} = \theta_{i,1} - 1$. The null of stationarity can now be expressed as the hypothesis that $\tilde{\theta}_{i,1} \neq 0$.[11] This is the Dickey–Fuller (DF) test for a unit root, OLS estimates of which are consistent provided the error term is white noise (e.g. Maddala and Wu, 1999; Baltagi, 2005). Adding lags of the dependent variable to ensure that the error term is serially uncorrelated gives the Augmented Dickey–Fuller (ADF) test. Observe also that $\tilde{\theta}_{i,1} = -1$ implies that the time span between periods t and $t-1$ is sufficiently long that past shocks have damped out entirely, in which case the period $t-1$ value for neighborhood economic status has no impact on the contemporaneous value.

A challenge in estimating (2.2) is that all of the panels used in this study are characterized by few time periods (small "T") and many cross-sections (large "N"). However, in recent years, there has been considerable effort devoted to panel unit root tests including situations with small T and large N (e.g. Harris and Tzavalis, 1999). In part, that is because the power of unit root tests is dramatically enhanced in a panel setting relative to that of a single time series (see Baltagi, 2005 for a discussion). I apply two of these methods here using routines available in Stata, Levin et al. (2002), and a Fisher-type method developed by Maddala and Wu (1999).[12]

The Levin et al. procedure is based on a multi-step process that constrains the autoregressive parameter to be alike across the different cross-sections and yields a corresponding test statistic, t^*. This test statistic has a standard normal distribution as T becomes large. The null in this case is that all of the time series in the panel contain unit roots, while the alternative is that none of them do. The Fisher test, in contrast, runs separate unit root tests on each of the panel's time series and forms a test statistic equal to

$$P = -2\sum_{i=1}^{N}\log(p_i), \tag{2.3}$$

where p_i is the p-value from the unit root test on cross-section i $(i = 1, \ldots, N)$. P is distributed Chi-square with $2N$ degrees of freedom as T becomes large and bootstrap methods have been used to establish critical values for finite samples (e.g. Maddala and Wu, 1999). In this case, the null is that all of the time series contain a unit root while the alternative is that at least one of them does not. The Fisher test is more general than the Levin et al. test as it allows for different autoregressive parameters across series and is preferred for that reason (see Baltagi, 2005). In the applications here, when running the Fisher test an ADF regression is used for the underlying unit root regressions. In addition, for both

[9] The 35 MSA balanced panel regression included county fixed effects in order to be more fully analogous to the Philadelphia County tract-level regression.

[10] When $\theta_{i,1}$ equals 1, $\log(y_{it})$ is sensitive to shocks far in the past, causing $\log(y_{it})$ to have an unstable mean and exploding variance as t becomes large.

[11] In the discussion below I emphasize Eq. (2.2), partly because that specification yields consistent estimates of $\tilde{\theta}_{i,1}$ even when $\theta_{i,1}$ equals 1, but also because Eq. (2.2) models *changes* in neighborhood relative income.

[12] The Levin et al. (2002) and Fisher routines are available as a user provided programs through Stata's website (http://www.stata.com).

the Fisher and the Levin et al. tests, one lag of the dependent variable is included in all of the individual time series to help eliminate serial correlation from the error terms.

A final issue concerns cross-sectional dependence. Both the Levin et al. and Fisher tests assume that the individual times series in the panel are independent and violation of that assumption can bias the estimated test statistics (e.g. Pesaran, 2003; Baltagi, 2005). To address that issue, I randomly sample one census tract from each city from among those tracts followed from 1950 to 2000. Because cities are spatially dispersed, and because tract relative income is based on income *relative* to its own metropolitan area, this helps to ensure cross-sectional independence. To allow for sampling variation, I sampled with replacement 1000 times and estimated the Levin et al. and Fisher tests for each of the 1000 samples. The distribution of the key test statistics based on the 1000 repetitions is portrayed in Figs. 1(a)–(b), and 2.

Figure 1(a)–(b) presents histograms of the Levin et al. (2002) autoregressive coefficient and its test statistics, t^*. As is apparent, the distribution of autoregressive parameters is tightly concentrated close to -0.5, well away from 0. The t^* values from the Levin et al. test are typically above 10. Moreover, the p-values associated with these estimates are all smaller than 0.000 with few exceptions.[13] Analogously, Fig. 2 presents the distribution of Fisher test statistics from the 1000 repetitions. These values are mostly above 300, and once again, the p-values associated with these test statistics are almost always smaller than 0.000. Together, these results overwhelmingly reject the null of a unit root. This provides further support for the idea that neighborhood economic status is a stationary process, and that neighborhoods cycle up and down in economic status around a stable long-run mean.[14]

[13] Because the p-values for the Levin et al. (2002) and Fisher tests are always so small they are not displayed to conserve space.

[14] I also ran a DF test on the Philadelphia Ward panel with a 50-year period length from 1900 to 2000. With T equal to 2 the only viable option was to constrain all of the coefficients in (2.2) to be alike across individual cross sections. The coefficient on the autoregressive term was -1.2 with a t-ratio of 6.2 based on OLS standard errors. Under the null of a unit root, the distribution of this test statistic is unknown owing to the very small value of T and the restrictive nature of the specification. Nevertheless, the high value of the autoregressive coefficient and the corresponding t-ratio are consistent with the results above.

3. Determinants of change in neighborhood economic status: model

The discussion thus far has established two stylized facts: (i) change in neighborhood economic status is common, and (ii) neighborhoods tend to cycle in economic status around a long-run mean over extended periods of time. This section models the determinants of these patterns. At the outset, it should be emphasized that not all factors that drive change in neighborhood economic status necessarily imply mean reversion and cycling. Instead, as suggested earlier, two broad mechanisms likely contribute to change in neighborhood economic status, aging housing stocks along with periodic redevelopment, and local externalities that arise from social interactions. As will become apparent, there are compelling reasons to anticipate that aging housing stocks contribute systematically to neighborhood economic cycles. However, whereas local externalities also drive change at the neighborhood level, most do not imply mean reversion, at least not directly.

To allow for these different sources of neighborhood change, expression (2.2) from the previous section is modified to include control variables for the age distribution of the housing stock and socioeconomic factors that proxy for social interactions. Each of these sets of covariates is discussed below, after which the empirical model is described.

3.1. Aging of the housing stock

Consider that cities tend to develop from the center outwards over time. Absent depreciation and redevelopment, the oldest dwellings would be found in the city centers, with the youngest structures on the urban/rural frontier. But this would lead to unusual sky-lines since, at the time of a city's initial development, land tends to be cheap and developers construct low-rise land-intensive buildings. The more familiar sky-lines of our larger cities owe their form to periodic waves of redevelopment.

Two factors drive the urban redevelopment cycle. First, from the hedonic house price literature, there is compelling evidence that most homes deteriorate with time, albeit at a slow rate that varies with the level of maintenance.[15] As homes depreciate, their flow of services declines and existing homes are eventually demol-

[15] Harding et al. (2007), for example, recently examined depreciation of housing capital controlling for the influence of maintenance and age related depreciation in a repeat sales context. They estimate the annual rate of depreciation for single family homes net of maintenance is

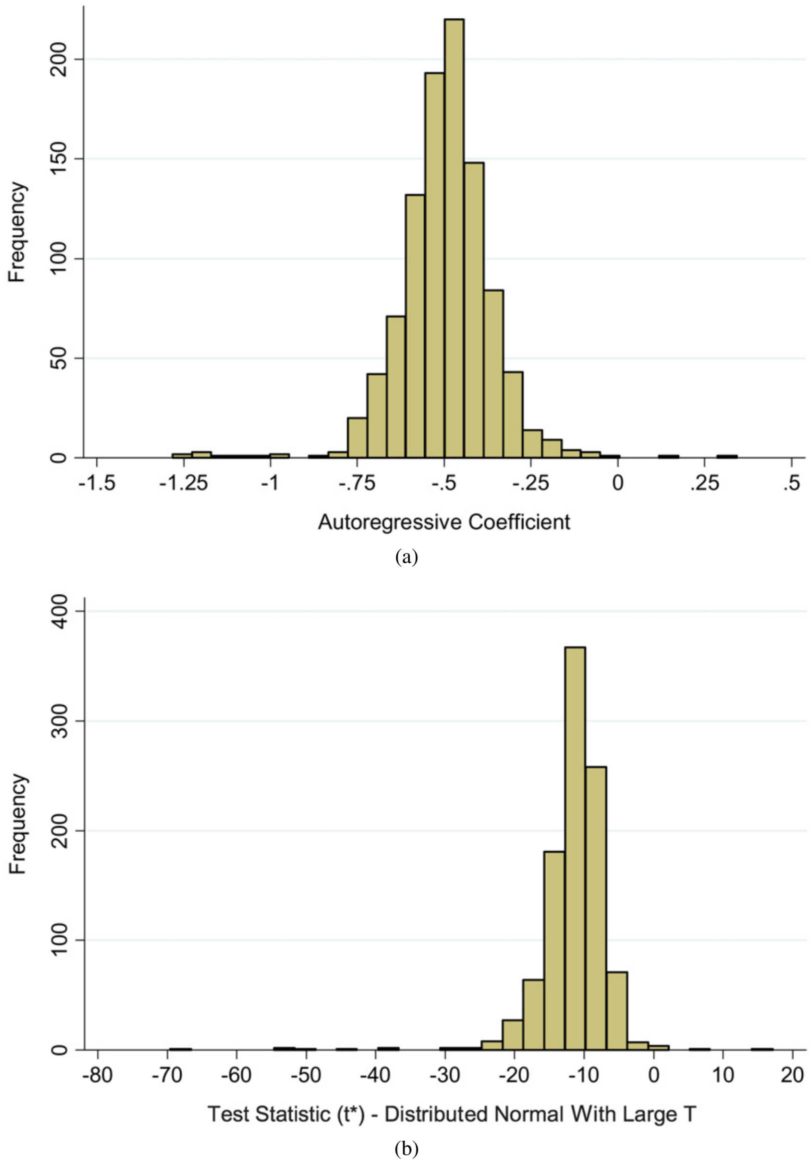

Fig. 1. Levin–Lin–Chu panel unit root test.

ished and the site redeveloped. This differs from economic "obsolescence." As time passes, existing housing capital is at risk of becoming economically obsolete even if it remains physically sound. This tends to occur in areas subject to sharp increases in real land prices that create pressure to increase the level of housing capital installed on the site. Under these conditions, physically sound buildings may be demolished and replaced with more expansive structures (see Rosenthal and Helsley, 1994 or Dye and McMillen, 2007, for example).

Because housing is a normal good, the slow decay (or increasing obsolescence) of existing housing stock encourages higher income families to move away from neighborhoods filled with older housing, ceteris paribus. Lower income families then take their place. This feature of urban housing markets is at the core of what has come to be referred to as the "filtering" model. The central role of filtering as the primary market source of low-income urban housing has been appreciated for many years. The seminal theoretical work is often attributed to Sweeney (1974), while Rothenberg et al. (1991) provides by far the most ambitious effort to ex-

roughly 2 percent per year. See also Smith (2004) for a further recent discussion of house price depreciation.

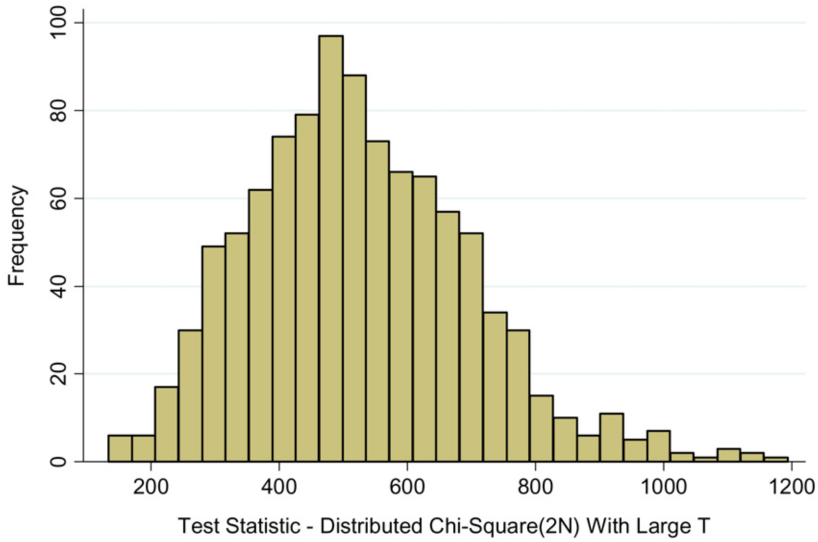

Fig. 2. Fisher panel unit root test.

amine filtering and related phenomena at an empirical level. Nevertheless, it should be emphasized that despite broad consensus that filtering is *the* private market source of low-income housing, there is a dearth of careful empirical work in this area.[16]

Several implications follow from these simple arguments. First, land within a given neighborhood tends to be developed in roughly the same time period, bearing in mind that time in this context is measured in decades. This suggests that there should be more variation in the age of the housing stock throughout an MSA than within individual neighborhoods. To confirm that premise, for each census tract in 2000, estimates were developed of the standard deviation of the age of the housing stock within the tract ($\sigma_{tract}^{HouseAge}$) as well as throughout the MSA ($\sigma_{MSA}^{HouseAge}$) to which the tract belongs.[17] The distribution of these measures across tracts and also the distribution of the difference between $\sigma_{tract}^{HouseAge}$ and $\sigma_{MSA}^{HouseAge}$ are reported in Table 4. At the median, $\sigma_{tract}^{HouseAge}$ is 3.65 years less than $\sigma_{MSA}^{HouseAge}$; at the 10th percentile, the differential is 9.47 years. In comparison, the average age of the US housing stock in 2000 was 33.09 years. These estimates confirm that

Table 4
Standard deviation of age of the housing stock in 2000 (in years)[a]

Percentile	Within individual census tracts	Within MSA to which a tract belongs	Difference between individual tract and the MSA to which it belongs
1	5.37	13.78	−13.70
10	9.87	17.06	−9.47
25	13.28	18.87	−6.69
50	16.99	21.02	−3.65
75	20.35	22.86	−0.79
90	22.98	22.86	1.38
99	25.99	23.78	4.67

[a] The average age of the housing stock in 2000 was 33.09 years.

[16] Additional important theoretical papers on filtering include Ohls (1975), Brueckner (1977, 1980), Bond and Coulson (1989), and Arnott and Braid (1997). Important empirical studies include Baer (1986), Weicher and Thibodeau (1988), and Aaronson (2001). Galster (1996) provides a nice review. Related work by Brueckner and Rosenthal (2006) considers the impact of aging of the housing stock on gentrification, a phenomenon especially relevant to growing older cities. Glaeser and Gyourko (2005), in contrast, emphasize that the presence of large stocks of existing housing can lead to precipitous declines in house prices in declining cities like Detroit relative to the impact of existing housing in growing cities. These studies suggest that the impact of aging housing stocks is likely not uniform across cities. In addition, Coulson and Bond (1990) question the degree to which age-related deterioration of the housing stock prompts filtering. They examine six cities in 1979 and 1980 and report low estimates of age-related effects on house prices. They then argue that the propensity of higher income families to seek out newer housing is driven by the tendency of newer homes to be larger. Harding et al. (2007), on the other hand, examine depreciation of housing capital for a national sample over the 1985 to 2001 period. Using repeat sales methods that control for house size and maintenance, their estimate of age-related depreciation (net of maintenance) is 1.94 percent per year, a value certainly high enough to support filtering over a period of decades.

[17] Estimates were formed using year 2000 because particularly detailed information is available on the age distribution of the housing stock in that year. Specifically, I controlled for housing age 0 to 4 years, 5 to 9, 10 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, and age 60 and over.

homes within a given neighborhood tend to be of more similar age relative to the homes throughout the neighborhood's broader MSA.

An implication of the relative similarity in age of nearby homes is that as homes within a higher-income neighborhood age, the neighborhood itself has a tendency to slowly filter down. That decline, however, is punctuated by increases in economic status as obsolete and dilapidated structures are demolished and replaced with new units attractive to higher income families. These ideas generate a set of predictions that will be tested. Specifically, newly built housing attracts higher income families and should elevate the future economic status of a neighborhood. "Middle-aged" housing should be associated with future decline in neighborhood economic status since these homes deteriorate with the passage of time (on average) but are not yet ripe for redevelopment. The presence of "old" housing, in contrast, is often a predictor of future demolitions, development of new housing, and gentrification. In this regard, old housing should be associated with an increase in the future status of the neighborhood. Evidence in support of these patterns would imply that filtering and redevelopment contribute to systematic cycles of neighborhood decline and renewal.

3.2. *Externalities*

Neighborhood externalities also contribute to migration and related change in neighborhood economic status. This likely occurs for two reasons: (i) certain types of families may behave in ways that generate *social capital and costs* for the neighborhood, influencing demand for that location, and (ii) families may choose to migrate into or out of a neighborhood based on the socio-demographic characteristics of their prospective neighbors. In this latter case, families are concerned with the *social status* of the neighborhood, but in both cases, social interactions lie at the core. Each of these mechanisms is now considered.

3.2.1. *Social capital and costs*

Three types of households seem especially likely to impart social capital on a neighborhood in a manner that would attract higher income families, and thereby, elevate the neighborhood's future economic status. These are the presence of prime aged workers, the presence of educated individuals, and the presence of homeowners. I briefly elaborate on each.

Prime aged workers likely bring financial resources to a neighborhood as compared to the elderly or the young. Such resources make it easier for families to maintain their homes, fund local public projects, and in general provide a positive resource for the community. The presence of educated individuals also brings resources into the neighborhood, except that in this case it is in the form of human capital as opposed to financial wealth, all else equal. It is well known, for example, that educated individuals commit fewer crimes and are more likely to be employed. Educated individuals also have skills that can potentially be used to help address challenges that may be faced by the community. For all of these reasons, children of educated individuals may be attractive playmates to parents deciding whether to move into or out of a neighborhood.

A recent literature has also argued that homeowners make better citizens (see Rhoe et al., 2000), for a review).[18] DiPasquale and Glaeser (1999), for example, provide evidence that homeowners are more likely to know the name of their congressional representatives, are more likely to belong to neighborhood groups, volunteer their time, garden, etc.[19] More generally, homeowners by definition are financially invested in their neighborhoods. As such, they have a greater incentive than renters to behave in ways that enhance the future economic status of the neighborhood, if not local property values. Partly for this reason, a host of public policy programs strongly advocate homeownership as a means of invigorating hard pressed urban neighborhoods.[20] Yet despite a long tradition in the United States of viewing homeownership as a public good, direct evidence that homeown-

[18] Green and White (1997) also argue that children of homeowners do better in school and have fewer teen pregnancies. Haurin et al. (2002) examine related issues.

[19] DiPasquale and Glaeser (1999) recognize that homeowners stay in their homes and neighborhoods longer than renters, and that length of stay could actually be the salient factor rather than homeownership. Indeed, the average length of stay in the home for homeowners is roughly ten years in the United States versus two years for renters (Rosenthal, 1988). However, when DiPasquale and Glaeser control for length of stay, they still find evidence that homeowners pay more attention to their local communities than do renters.

[20] Cummings et al. (2002) describe a number of such programs as part of their study of the effects of Philadelphia Nehemiah programs designed to promote homeownership in the center of the most hard-pressed neighborhoods. Cummings et al. note that the City of Philadelphia has "... long encouraged homeownership as part of its overall community development strategy ..." Further, a primary goal stated in the strategic plan of the Office of Housing and Community Development (OHCD) of the City of Philadelphia is "promoting homeownership and housing preservation. ... to more effectively support economic development and reinvestment in Philadelphia, the City will continue to emphasize homeownership and preservation of the existing occupied housing stock" (OHCD, 1997, p. 9; Cummings et al., 2002, p. 332).

ership produces spillover effects has been quite limited.

On the negative side, crime clearly imposes a social cost on a community. To the extent that certain types of families commit more crimes, the presence of such households will discourage investment in the community. That, in turn, will lower property values, attract lower income households, and reduce the economic status of the neighborhood. Recent papers by DiPasquale and Glaeser (1999) and Glaeser and Sacerdote (1999) suggest that cities tend to be subject to higher crime rates in part because criminals are less likely to be apprehended in populous areas. More generally, this implies that densely developed neighborhoods may be more subject to crime and, therefore, suffer lower expected future economic status, all else equal.

To allow for each of these factors, in the regression models to follow, controls are provided for the density of development in the neighborhood, homeownership rate, the distribution of education (percent less than high school, high school degree, some college, college degree or more), and the age distribution of the population (percent age 15 or less, age 16 to 29, age 30 to 54, and age 55 and over).

Also included in the model is the neighborhood income inter-quartile range. For any given average level of income in the neighborhood, greater variation in neighborhood income implies the presence of higher and lower income individuals. The presence of higher income individuals may generate social capital that could elevate a neighborhood's future economic status, but the presence of lower income families could do the opposite. Theory, therefore, is ambiguous about the effect of income mixing on a neighborhood's future status.

3.2.2. Social status

If families choose to migrate into or out of a neighborhood because they care about a neighborhood's social status, this could further affect the future economic standing of the neighborhood. A prominent example of this concerns the racial composition of the neighborhood. "White flight" was first used to describe the huge numbers of white central city households who moved to the suburbs following the race riots of the 1960s.[21] Implicit in the phrase is the idea that white families do not want to live in close proximity to African Americans. Because minorities tend to be of lower economic status than whites, sorting by race and ethnicity has an indirect effect on neighborhood economic status. To allow for such effects, in the regressions to follow, controls are provided for the racial/ethnic composition of the neighborhood, specifically, the percent of the neighborhood's population that is African American and the percent that is Hispanic.[22]

Another instance when families may care about the attributes of their neighbors concerns the presence of place-based subsidized rental housing. Federal law, for instance, mandates that public housing projects serve primarily low-income families. At its peak, there were roughly 1.3 million public housing units in the US in 1990. New construction of public housing was largely stopped in the mid-1980s, but in 1987 the Federal government began subsidizing construction of lower-moderate income housing through the Low Income Housing Tax Credit (LIHTC). Since 1987, roughly 1.5 million housing units have been built or substantially refurbished through the LIHTC program. It is possible that NIMBY-type behavior ("not in my back yard") could cause higher income households to shy away from neighborhoods in which such housing is situated.[23]

3.3. Empirical model

To evaluate the determinants of change in neighborhood economic status I modify the unit root regression in Eq. (2.1) by adding controls for the house age (*HouseAge*) and SES (*SES*) variables described above. Also included in the model is the distance measured in miles to the census tract with the highest population density in year 2000 (*Distance*). The distance variable helps to control for correlation between the location and timing of development. To maintain focus on the change in neighborhood economic status, the dependent variable is expressed as the log change in a community's economic status between periods. The initial model specification is:

[21] This process is often thought to be particularly sensitive to "tipping" points as discussed by Schelling (1971) and others. Accordingly, earlier versions of this paper allowed for non-linear effects from neighborhood racial composition. Although evidence of non-linearities was present, those effects had little influence on the main results. For that reason, and because tipping points are not central to this paper, a linear specification is adopted below.

[22] In the 1950 and 1960 data, Hispanics are defined as individuals whose country of origin was Mexico.

[23] See Olsen (2003) and Eriksen and Rosenthal (2007) for further discussion of the public housing and LIHTC programs. Also, as noted by Cummings and DiPasquale (1999), many LIHTC units are in fact of high quality compared to other low-income housing. For that reason, it is possible that LIHTC development could actually improve the appeal of lower-income neighborhoods while reducing the appeal of higher income communities. Eriksen and Rosenthal (2007) report evidence consistent with such effects.

$$\log(y_{i,t}/y_{i,t-1}) = \theta_i + b_1 X_{i,t} + \theta_1 \log(y_{i,t-1}) + e_{i,t}. \tag{3.1}$$

where i denotes the census tract, t is the current decade, and X_{it} denotes the *HouseAge*, *SES*, and *Distance* controls. The precise elements of *HouseAge* and *SES* are described later in the paper.

In principle one could estimate (3.1) in first differences as with the Arellano–Bond estimator (see Cameron and Trivedi, 2005 and Baltagi, 2005 for discussions). This would have the appeal of differencing away all of the time-invariant tract specific effects. In practice, however, for several reasons I do not focus on such a model. First, time-differenced estimators impose the restriction that the coefficients on the control measures do not change between periods. Especially for the SES variables that is unappealing given that social norms change over time, and certainly between 1950 and 2000. Second, the level of detail about the age distribution of the housing stock is greater in later periods, and that creates some incentive to focus on more recent decades. Third, and most important, all of the contemporaneous house age and SES variables are potentially endogenous. That is because households and developers choose a neighborhood in which to invest their time and resources. As such, I cannot rule out the possibility that the current mix of households and housing stock reflect the influence of unobserved neighborhood attributes that are embedded in the model error term.[24] This complicates identification by greatly increasing the number of potentially endogenous variables while also precluding the use of contemporaneous values for *HouseAge* and *SES* as instruments.

As an alternative to (3.1), I focus on a modification of expression (2.2) and estimate the following model:

$$\begin{aligned}\log(y_{i,2000}/y_{i,1990}) &= \theta_{MSA} + b_1 HouseAge_{i,2000-k} \\ &+ b_2 SES_{i,2000-k} + b_3 Distance_i + \tilde{\theta}_1 \log(y_{i,2000-k}) \\ &+ \log(y_{i,2000-k}/y_{i,2000-k-10}) + e_{i,2000}.\end{aligned} \tag{3.2}$$

In this expression, k is the number of decades that time-varying covariates are lagged. As will become apparent, k is varied across specifications in a manner that helps to ensure that the regressors are exogenous and also to illuminate differences in the impact of the *HouseAge* and *SES* control measures.

A key feature of (3.2) is that the dependent variable is restricted to the log change in economic status between 1990 and 2000. Expression (3.2) is then estimated as a single cross section including the lag level of neighborhood economic status, $\log(y_{i,2000-k})$, to control for mean reversion, as well as one lag of the dependent variable, $\log(y_{i,2000-k}/y_{i,2000-k-10})$, to control for serial correlation in the model error term. This structure mimics that of (2.2) except that (3.2) controls for MSA rather than tract specific fixed effects, θ_{MSA}.[25] Identification, therefore, is based on within MSA variation as opposed to tract-specific variation over time.

As an initial attempt to estimate (3.2), I set k equal to 1 decade indicating that 1990 values are used for the control measures. Given that the dependent variable is the difference between year 2000 and year 1990 neighborhood economic status, it should be emphasized that the 1990 covariates are best interpreted as contemporaneous values and are likely endogenous. For that reason, (3.2) was estimated by two-step GMM treating the 1990 economic status of the neighborhood, *HouseAge*, and *SES* variables as endogenous. For instruments I used 1980 and 1970 house age and SES factors, as well as the log levels of the neighborhood's economic status in those years. This approach hinges on the assumption that the 1980 and 1970 neighborhood attributes are exogenous to the community's change in economic status in the 1990s. As an approximation, this seems reasonable as the alternative would be that households and developers in the 1970s made decisions based on anticipated changes in neighborhood economic status in the 1990s. Although I cannot rule out that possibility, it seems likely to be of second order concern.[26]

Diagnostic tests on the instruments are provided in Table 5 where results for five different models are presented. The first set of results corresponds to the full set of census tracts in the sample. The remaining four sets of results correspond to census tracts belonging to the different economic quartiles in 1990 based on relative income within individual MSAs in a manner analogous to that of Table 1 (bottom quartile, 2nd quartile, 3rd quartile, and top quartile). In the full sample case

[24] For example, if a 10-year development plan is ongoing, households and developers may choose to invest in the community in anticipation of the rising status of the neighborhood. This could cause the house age and SES variables to be correlated with the model error term.

[25] This includes the city-wide level of income, the extent of racial segregation, fiscal policy, and broader macroeconomic conditions throughout an MSA.

[26] On this point, it is also worth noting that few families remain in their homes and neighborhoods longer than ten years. For this reason, lagged SES control measures reflect the attributes of a completely different population from that in the neighborhood in 1990. This ensures that the influence of lagged SES variables is channeled through dynamic spillover effects as opposed to a change in the economic status of an existing set of residents.

Table 5
Diagnostic tests for GMM estimates with 1990 covariates using 1980 and 1970 house age, SES, and neighborhood economic status as instruments

	Full sample	Tract economic status relative to MSA in 1990			
		1st quartile	2nd quartile	3rd quartile	4th quartile
Number of Endog Vars	18	18	18	18	18
Number of instruments	31	31	31	31	31
Under-identification test[a]	44.2	32.76	27.98	17.18	12.46
p-value	0.00	0.01	0.02	0.31	0.64
Weak instrument test[b]	1.38	1.02	0.87	0.54	0.39
p-value	NA	NA	NA	NA	NA
Over-identification test[c]	111.92	21.71	37.23	38.78	24.50
p-value	0.00	0.00	0.00	0.00	0.00
Observations	43,422	10,579	10,622	11,090	11,126
MSA fixed effects	270	270	268	267	269

[a] The under-identification test is the Anderson Canonical Correlations LR Test where the null hypothesis is the rank of is the matrix of instruments is one less the number of endogenous variables. The test is distributed Chi-squared with 14 degrees of freedom.

[b] The weak instrument test is the Stock and Yogo (2005) modified version of the Cragg and Donald (1993) Test Statistic. The critical values with 17 endogenous variables are unknown. In models with three or fewer endogenous variables a test statistic greater than 10 often indicates less than 10 percent IV bias.

[c] The over-identification test is the J Hansen (1982) test of the over identifying restrictions of the excluded instruments. The joint null hypothesis is the instruments are uncorrelated with the error term and correctly excluded from the estimated equation. Rejection of the null hypothesis either indicates invalid instruments or model misspecification. The test statistic is distributed Chi-squared with 13 degrees of freedom.

and also the first two quartile samples the models pass the under-identification test. However, for the upper two quartile samples the models fail that test. More generally, the first-stage F-statistics on the instruments are sufficiently low that all of the models likely suffer from weak instrument bias.[27] In addition, the models all fail the Hansen-J test of the overidentifying restrictions. Assuming that the 1980 and 1970 neighborhood attributes are exogenous, this latter test suggests that lagged house age and SES factors have a meaningful role in the structural model and should not be excluded: in effect, history matters.

Given the diagnostic tests in Table 5, I abandon the GMM approach and proceed as follows. It seems plausible that 1980 neighborhood attributes are exogenous to change in a community's economic status between 1990 and 2000. Accordingly, in most of the specifications to follow, I estimate (3.2) setting all of the covariates to their 1980 or earlier values ($k > 1$) in a manner to be clarified. All of the estimation is then carried out using ordinary least squares.

[27] The largest weak instrument test statistic is 1.38 for the full sample. In comparison, recent simulations by Stock and Yogo (2005) and Staiger and Stock (1997) based on three or fewer endogenous variables suggest that IV estimates display up to 10 percent bias when the relevant F-statistic per endogenous variable is below 10 and substantial bias when that test statistic is below 5. Although the literature has yet to establish critical values for weak instrument bias when many endogenous variables present, the test statistics in Table 5 are nevertheless quite low. See Murray (2006) and Cameron and Trivedi (2005) for recent discussions of these issues.

4. Determinants of change in neighborhood economic status: evidence

This section presents estimates from several variants of expression (3.2) and seeks to provide evidence on two issues. First, that filtering and externalities arising from social dynamics are fundamentally different processes that affect change in neighborhood economic status in very different ways. Second, the section examines the impact of individual control measures (e.g. neighborhood house age distribution, presence of homeownership, minorities, etc.) on change in a community's economic status. I begin with the nature of filtering versus spillover effects arising from social dynamics.

4.1. Filtering versus externalities from social dynamics

In Table 6, OLS estimates of model (3.2) are presented for two specifications. The first specification includes all of the covariates described by expression (3.2), including both the *HouseAge* and *SES* variables. The second specification omits the SES variables. In both cases, the omitted house age category includes homes age 40 or older. In addition, k is set equal to 2 decades in both cases indicating that all of the covariates are based on 1980 values. Finally, each of these models is estimated twice: first for the entire set of MSAs in the sample, and then again for the 35 MSAs highlighted earlier. To simplify review, only the house-age coefficients are reported.

Table 6
Independence of filtering and SES effects dependent variable: log y_{2000}/y_{1990}.
(Absolute value of t-ratios based on robust standard errors in parentheses)

	35 MSAs		All MSAs	
	Exclude SES[a]	Include SES[a]	Exclude SES[a]	Include SES[a]
% Homes 0–9 yrs	−0.0827	−0.0577	−0.0721	−0.0588
	(6.30)	(3.76)	(12.34)	(8.32)
% Homes 10–19 yrs	−0.1352	−0.0892	−0.1150	−0.0885
	(9.13)	(5.58)	(15.71)	(11.43)
% Homes 20–29 yrs	−0.0724	0.0080	−0.0619	−0.0272
	(5.29)	(0.53)	(8.36)	(3.41)
% Homes 30–39 yrs	−0.0500	0.0091	−0.0263	−0.0067
	(2.11)	(0.37)	(2.02)	(0.50)
Observations	9419	9419	43,654	43,654
MSA fixed effects	34	34	271	271
Adj. R-squared	0.03	0.08	0.03	0.05

[a] SES controls are those reported in Table 8. The omitted house age category is the percent of homes 40 years and older. Variables other than the house age distribution are not reported to conserve space. All control variables are based on 1980 values.

Table 7
Relative significance of lagged house age and SES control measures on change in neighborhood economic status between 1990 and 2000.
Dependent variable: $\log(y_{2000}/y_{1990})$[a]

	35 MSAs with covariates from					All MSAs with covariates from		
	1990	1980	1970	1960	1950	1990	1980	1970
$F(4, \text{Obs})$ for *House Age*[b]	13.75	14.14	–	–	–	77.75	50.63	–
$F(3, \text{Obs})$ for *House Age*[b]	–	–	21.07	10.95	18.31	–	–	76.69
$F(13, \text{Obs})$ for *SES*[b]	113.58	42.42	24.32	19.13	15.06	437.24	98.68	58.61
$F_{SES}/F_{House\ Age}$	8.26	3.00	1.15	1.75	0.82	5.62	1.95	0.76
Observations	9419	9419	9402	9406	9419	48,622	43,654	43,397
MSA fixed effects	34	34	34	34	34	325	271	270
R-squared (within)	0.19	0.09	0.07	0.06	0.06	0.16	0.06	0.05

[a] The complete set of results for these regressions are provided in the appendix, Table A.4. All specifications include *Distance*, the 1990 log level of neighborhood relative income, and one lag of the dependent variable to control for serial correlation.

[b] For 1990 and 1980 covariates, house age variables are as in Table 8 and include the percent of homes 0 to 9 years, 10 to 19 years, 20 to 29 years, and 30 to 39 years with 40+ years as the omitted category. For 1970, 1960, and 1950, the omitted house age category is the percent of home age 30+ years. For all years SES controls are the same as for Table 8 with the exception of 1990 for which the percentage of home produced through the Low-Income Housing Tax Credit (LIHTC) program are included.

Comparing estimates across the two models, it is clear that omitting the *SES* measures has little effect on the qualitative nature of the house age coefficients, and a relatively modest impact on the magnitudes of those estimates. Regardless of which sample is used and whether the SES controls are included in the regression, the coefficient on newly built housing (under 10 years in age) is close to −0.07. That coefficient increases by 50 to 100 percent for homes age 10 to 19 years, and then diminishes sharply for older housing categories (age 20 to 29 and 30 to 39). Overall, these patterns suggest that the influence of filtering on change in neighborhood economic status, while not strictly independent of the *SES* factors, is partly so. That result is consistent with the premise that filtering and neighborhood externalities are quite different processes.

To explore this idea further, Table 7 presents summary measures of the joint significance of the *HouseAge* and *SES* control measures for a series of regressions that specify increasingly deeper lags of those variables. In the first five columns, results are presented for the 35 MSA census tract panel with the *HouseAge* and *SES* variables set to 1990 values, 1980 values, and on back to 1950 values, respectively. In the last remaining three columns of the table, an analogous set of regressions are presented for the complete sample of MSAs with the *HouseAge* and *SES* variables set to 1990, 1980 and 1970 values, respectively. For those models in which 1990 values are used for the house age and SES covariates, k was also set to 1 decade for $\log(y_{i,2000-k})$ and $\log(y_{i,2000-k}/y_{i,2000-k-10})$. In all other cases the lagged economic status variables were based on 1980 values with k set to 2. Finally, to focus on the joint significance of the *HouseAge* and *SES* variables, only the F-statistics on the *HouseAge* and *SES* variables were presented in the table. These statistics test for whether the house age and SES variables are jointly equal to zero. The full set of coefficient estimates from these regressions is provided in the appendix (Table A.4).

Several patterns are noteworthy in Table 7. First, regardless of the sample used, when the covariates are set to 1990 values the F-statistic on the SES factors is several times larger than the F-statistic for the house age variables. To some extent this reflects that there are 13 SES variables and only 4 house age variables. Nevertheless, the relative magnitude of the two F-statistics strongly suggest that as a group, near-term SES factors explain much more of the change in neighborhood economic status as compared to the influence of the near-term house age distribution.

Comparing F-statistics across columns, it is striking that the *HouseAge* F-statistics largely do not diminish with deeper lags back to 1970. In contrast, the *SES* F-statistics decline sharply with deeper lags, including not only the shift from 1990 to 1980 controls, but also the shift from 1980 to 1970 controls. Given these patterns, when using 1960 and 1950 covariates, the F-statistics for the *HouseAge* and *SES* factors are similar in magnitude.

Overall, the patterns in Tables 5 and 6 provide evidence that filtering and externalities associated with social dynamics are very different processes and that both contribute to change in a neighborhood's future economic status. SES factors have the greater impact in the short run (e.g. one decade), but the effect of SES attributes declines sharply with time. In contrast, the influence of the age distribution of the housing stock is remarkably persistent even over a fifty year horizon. As a consequence, the impact of deeply lagged house age and SES factors on future changes in a neighborhood's economic status is similar in magnitude.

4.2. House age and SES covariate effects

We turn now to the impact of the individual house age and SES covariates on change in neighborhood economic status in the 1990s. Moreover, to allow for greater heterogeneity, the census tract samples are once again stratified into economic quartiles based on relative income within the individual MSAs in 1990. As a reminder, in all cases k is set equal to 2 decades indicating that all of the time-varying covariates are measured as of 1980.

Consider first the *HouseAge* variables in Table 8 for both the "Full Sample" and also for the different sample quartiles. Plots of these variable coefficients are also provided in Fig. 3 with age of the housing stock on the horizontal axis. For the full sample and also the top three income quartiles there is a roughly U-shaped pattern with the trough centered on housing aged 10 to 19 years in 1980. This indicates that the presence of both newly built housing and older housing in 1980 enhance the economic status of a neighborhood in the 1990s relative to the presence of "middle-aged" housing. Observe also that housing age 40 years or older has the most positive impact on the change in a neighborhood's economic status. This later feature is also evident for the lowest income quartile sample, but for that group the pattern of coefficients is largely flat for housing age 0 to 39 and then higher for older housing.

There are perhaps two possible interpretations of these patterns. The first is that high quality homes tend to be preserved and for that reason, communities with filled with old homes have high quality housing and are sought out by higher income families. For several reasons, however, this does not seem like a very credible explanation for the patterns in Fig. 3. First, the argument is focused on the level of a neighborhood's economic status rather than change in the community's economic standing. Second, the presence of high-quality older (age 40 or more) housing would not account for why housing age 10 to 19 has a more negative coefficient as compared to housing age 20 to 29 or 30 to 39 as occurs for the upper three quartile samples. Third, housing is durable and the great majority of homes likely are retained beyond age 40. For that reason, housing over age 40 is almost certainly not of higher quality, on average relative to younger portions of the housing stock. Consistent with that view, as noted earlier, numerous hedonic housing price studies in the literature find that housing tends to deteriorate with age.

A more likely explanation for the patterns in Fig. 3 is that housing age 40 or more tends to be of unusually low quality owing to deterioration over time. Moreover, sufficiently old housing is increasingly at risk of redevelopment. The presence of older housing in 1980, therefore, likely contributed to gentrification in the 1990s as higher income families occupied newly rebuilt and/or refurbished neighborhoods (see Brueckner and Rosenthal, 2006 for further evidence on this point). Under this view, middle aged housing from 1980 decays with the further passage of time, but is not yet old enough to warrant demolition in the 1990s. Middle aged housing in 1980, therefore, encourages subsequent in-movement of lower income families. In short, the patterns in Fig. 3

Table 8
Tract change in economic status in the 1990s by income quartile with 1980 covariates.
Dependent variable: $\log(y_{2000}/y_{1990})$

1980 Covariates	Full sample	Tract economic status relative to MSA in 1990			
		1st quartile	2nd quartile	3rd quartile	4th quartile
% Homes 0 to 9 yrs	−0.0588	−0.058	−0.0325	−0.0935	−0.1136
	(8.32)	(3.09)	(2.58)	(7.21)	(7.63)
% Homes 10 to 19 yrs	−0.0885	−0.0524	−0.09	−0.1448	−0.15
	(11.43)	(2.26)	(6.34)	(10.81)	(9.34)
% Homes 20 to 29 yrs	−0.0272	−0.0206	−0.0614	−0.1061	−0.0614
	(3.41)	(0.90)	(4.48)	(7.63)	(3.33)
% Homes 30 to 39 yrs	−0.0067	−0.0688	−0.0443	−0.0017	0.0369
	(0.50)	(2.33)	(2.00)	(0.07)	(1.09)
% Public housing units	0.0007	−0.0005	−0.074	−0.0198	−0.0429
	(0.07)	(0.03)	(2.10)	(0.98)	(1.78)
Density (1000 units/mile)	−0.0045	−0.0062	−0.0046	−0.005	−0.0006
	(13.21)	(8.05)	(7.12)	(6.49)	(0.95)
Homeownership rate	0.064	0.0357	0.1231	0.0811	0.0285
	(8.51)	(1.93)	(8.82)	(5.58)	(1.69)
% High school degree (heads)	−0.0467	0.0074	−0.1045	−0.0842	−0.0177
	(3.18)	(0.22)	(3.74)	(2.77)	(0.43)
% College degree+ (heads)	0.2557	0.4628	0.2706	0.1934	0.0741
	(18.03)	(10.11)	(8.82)	(7.17)	(2.30)
% Married (men 18+)	−0.2528	−0.2661	−0.133	−0.1232	−0.1569
	(14.50)	(7.14)	(3.98)	(3.57)	(3.87)
% Pop. < age 15	0.0956	−0.0807	−0.3804	−0.0497	0.1870
	(2.38)	(0.88)	(4.83)	(0.64)	(1.98)
% Pop. age 15 to 29	−0.3414	−0.5066	−0.4503	−0.2772	−0.2877
	(13.06)	(7.88)	(8.31)	(5.27)	(5.32)
% Pop. age 55+	−0.0653	−0.1325	−0.2792	−0.155	−0.1502
	(2.47)	(1.97)	(5.49)	(3.24)	(2.72)
% African American (heads)	−0.0566	−0.0348	−0.0739	−0.1081	−0.253
	(9.53)	(2.99)	(7.17)	(8.28)	(10.72)
% Hispanic (heads)	−0.005	0.0383	−0.0084	−0.0912	−0.0784
	(0.47)	(1.86)	(0.41)	(3.29)	(1.58)
IQR Log Inc (year 2000 $)	0.0000	−0.0012	0.0004	0.0002	0.0000
	(0.05)	(2.70)	(1.35)	(0.65)	(0.06)
Distance to 2000 CBD (miles)	0.0012	0.0009	0.001	0.0011	0.0015
	(11.39)	(3.83)	(6.16)	(6.12)	(5.55)
$\log(Y_{1980})$	−0.1732	−0.2177	−0.086	−0.0379	−0.0581
	(20.21)	(9.48)	(4.25)	(1.99)	(4.03)
$\log(Y_{1980}) - \log(Y_{1970})$	0.0791	0.1145	0.0918	0.0828	0.0344
	(12.24)	(6.61)	(6.91)	(6.89)	(3.04)
Observations	43,654	10,619	10,661	11,124	11,250
MSA fixed effects	271	270	271	271	269
R-squared (within)	0.06	0.07	0.06	0.06	0.04

t-ratios based on robust standard errors in parentheses.

are consistent with the filtering/redevelopment view of housing markets, and also with the presence of long-run cyclical movements in neighborhood economic status.

Results from the SES factors are also of interest and are displayed both in Table 8 and in Fig. 4. Observe first that density, measured as 1000 housing units per square mile, has a negative impact on the future economic status of neighborhoods. In addition, that effect declines roughly monotonically with an increase in the neighborhood's economic status. These patterns are consistent with the idea that for most neighborhoods, high density fosters negative spillovers, but higher income communities are able to shelter themselves from the negative effects of high density development.

The presence of homeowners in 1980 elevates the economic status of the neighborhood for all communities in the 1990s, but most for the middle-income neighborhoods. The coefficients on homeownership for the second and third quartile samples are 0.123 and 0.081, respectively. This indicates that a 10 percentage point

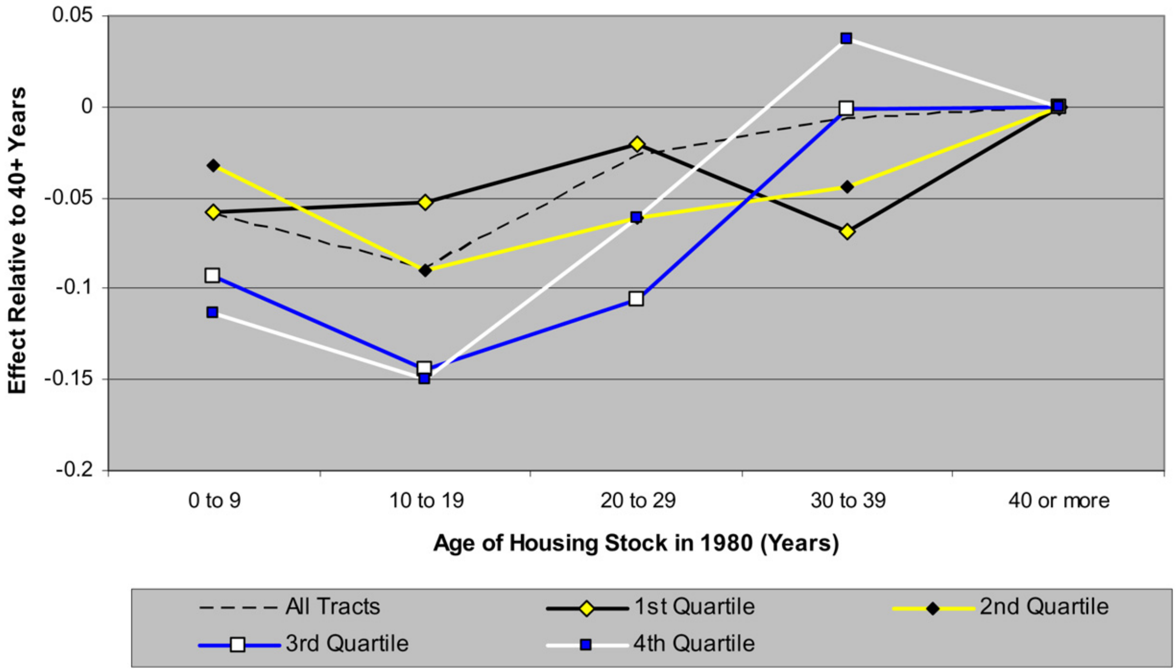

Fig. 3. 1980 age of housing stock and neighborhood change between 1990 and 2000.

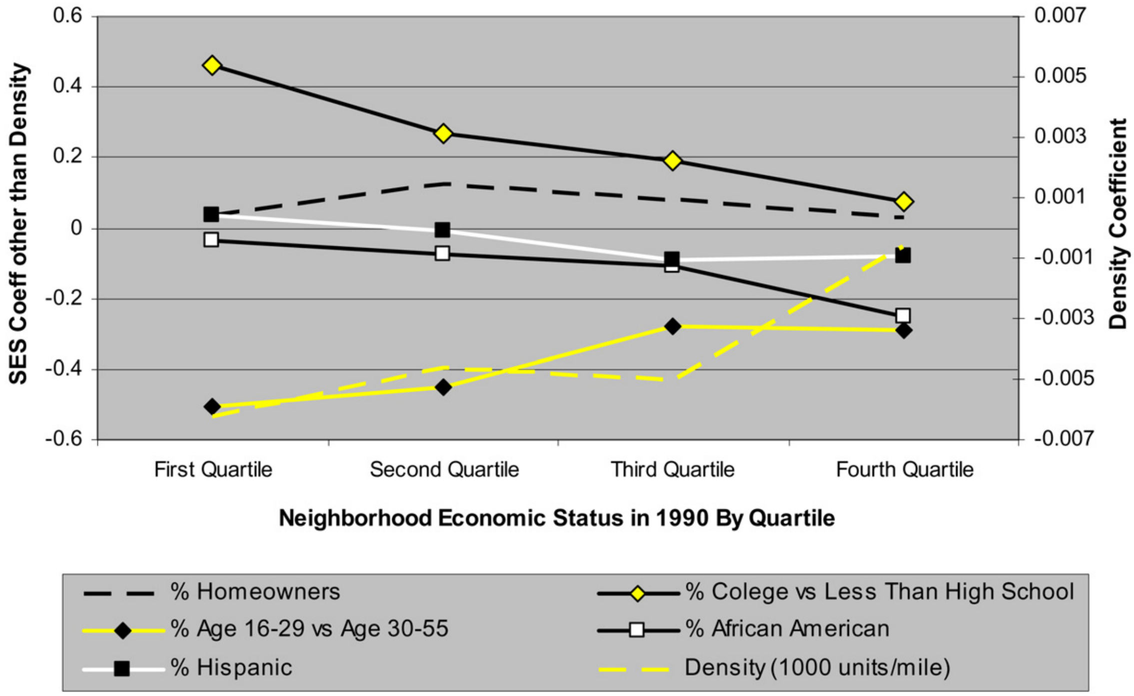

Fig. 4. Impact of 1980 socio-demographic factors.

higher homeownership rate in 1980 would boost the economic status of second- and third-quartile communities by 1.23 and 0.81 percentage points in the 1990s. Among low and high income neighborhoods the corresponding effect is roughly 0.3 percentage points. These findings are consistent with the idea that homeowners behave in ways that enhance their property values, and the future economic status of their neighborhoods.

These patterns also lend support to public policies that seek to promote homeownership as a means to invigorate low-moderate income communities given the large coefficient for the second quartile sample.

Analogous results suggest that the presence of college educated individuals brings human capital to a community that also foster positive spillover effects. In Table 8 and Fig. 4, observe that the presence of college educated individuals has a positive impact on the future economic status of a neighborhood. Moreover, that effect is pronounced among low-income communities and declines monotonically in magnitude with the economic status of the neighborhood. For first-quartile neighborhoods, a 10 percentage point increase in the 1980 level of college degree individuals would cause the neighborhood's economic status to rise in the 1990s by 4.63 percentage points; the corresponding effect for fourth-quartile communities is 0.74 percentage points. Once again, there is a tendency for lower-income areas to benefit most from individuals that provide social capital.

The same qualitative pattern is evident with regard to the presence of prime age individuals, defined here as individuals age 30 to 55 in 1980. With this group as the omitted age category in the regressions, the coefficients on the presence of individuals age 15 to 29 in 1980 are plotted in Fig. 4. Relative to prime age workers, the presence of individuals age 15 to 29 has a sharply negative impact on change in neighborhood economic status in the 1990s that is most pronounced for the lowest income communities. For the first quartile sample, the coefficient is −0.51 while for the fourth quartile sample the coefficient is −0.29. The negative sign on these coefficients is consistent with the idea that prime age individuals bring greater financial and job-network resources to a community. The larger coefficients for the lower-income quartiles suggest once more that these areas benefit most from related spillovers.

Finally, the influence of race is also striking. A 10 percentage point increase in the 1980 African American presence would decrease a first-quartile neighborhood's economic status in the 1990s by 0.35 percentage points. In contrast, the corresponding effect on fourth-quartile communities is 2.5 percentage points. For Hispanics, the analogous values are positive 0.38 percentage points for first-quartile neighborhoods and negative 0.78 percentage points for fourth-quartile locales. These and other race-related patterns in Table 8 allow for two broad characterizations: first, the presence of minorities in 1980 generally had a negative impact on change in neighborhood economic status in the 1990s, and second, that adverse effect increased with the economic status of the community.

It is difficult, of course, to know exactly what these race-related results reflect. African American and Hispanic households are known to have less wealth and may therefore bring less social capital to a community. But that would not explain the disproportionate effect of minority presence on higher income communities. An alternative explanation is that lower income minorities seek out communities already populated with such families, and/or that higher income white families flee neighborhoods where minorities are present. Both would contribute to the downward slope in the race-related plots in Fig. 4.[28]

5. Conclusions

This paper began with an observation: most low-income urban families in the US occupy old homes built originally for higher income households. That simple observation implies regular but long-running cycles of neighborhood decline and renewal as homes age and are replaced. Superimposed on those patterns are the shorter run effects of a somewhat independent process, neighborhood externalities. This occurs when certain types of individuals behave in ways that generate social capital or costs, affecting demand for the neighborhood, and when individuals base migration decisions on the attributes of their prospective neighbors. To explore these patterns, I assembled a unique set of panels that follow neighborhoods on a geographically consistent basis over extended periods of time.

Findings indicate that change in neighborhood economic status is common over a sufficiently long time horizon. Roughly two-thirds of urban neighborhoods in 1950, for example, were of quite different economic status in 2000; on average, urban neighborhoods shift 13 percent in economic status with each decade. Panel unit root tests based on a balanced panel of roughly 9500 census tracts (one seventh of the United States) followed from 1950 to 2000 support the idea that neighborhoods cycle up and down over extended periods of time. In Philadelphia County, an area followed from 1900 to 2000, a complete cycle appears to last up to 100 years. Thus, most urban neighborhoods are not static

[28] The remaining coefficients in Table 8 mostly are insignificant (e.g. the public housing and income inter-quartile variables) or reflect patterns discussed earlier (e.g. the lagged neighborhood economic status). Observe also that tracts further from the center of town (*Distance*) exhibit greater increase in economic status in the 1990s. This is consistent with development of higher income suburbs during this period.

but will be of fundamentally different economic status twenty to forty years in the future.

Evidence also confirms that filtering and neighborhood externalities both contribute to change in neighborhood economic status but in very different ways. Whereas externalities arising from a neighborhood's socio-demographic composition are most pronounced in the short run (i.e. a decade), the influence of the age distribution of the housing stock is very persistent over several decades. In addition, it is clear that relative to both new and older housing, the presence of middle aged homes reduces the degree to which a community may rise up in economic status twenty years ahead. This is consistent with the idea that middle aged housing deteriorates with the further passage of time, but is not yet ripe for imminent demolition and replacement. In contrast, the presence of old housing is a forerunner to urban redevelopment and gentrification as higher income families are attracted to newly redeveloped and refurbished neighborhoods. Overall, these patterns confirm that filtering and local externalities affect a neighborhood's future economic status but through quite different mechanisms.

Many of the estimated effects of socioeconomic factors are also economically important. For example, a 10 percentage point higher homeownership rate in a second quartile income neighborhood in 1980 would cause the neighborhood's economic status to rise by 1.23 percentage points in the 1990s, all else equal. This is consistent with the idea that homeowners provide social capital and lends support to recent aggressive efforts by Federal and local governments to promote homeownership in low-moderate income communities. Moreover, the presence of homeowners, college educated individuals, and prime age workers, all appear to have more pronounced positive effects on lower income communities. This is suggestive that low income areas benefit most from individuals who provide social capital and that attracting such individuals may be good public policy if the goal is to elevate a given neighborhood.

As a final perspective, it is important to recognize that policies designed to raise a neighborhood's economic status do not necessarily alleviate poverty. Instead, evidence presented here and in Brueckner and Rosenthal (2006) suggest that such efforts could simply force existing low-income residents to relocate to other parts of a city. This issue recently received considerable public press when President Clinton located his office in Harlem.[29] Although important, this issue is left for further research.

Acknowledgments

Funding for this project from the John D. and Catherine T. MacArthur Foundation, the Ford Foundation, and the Syracuse University Center for Policy Research is gratefully acknowledged. I am also grateful for helpful comments from Jan Brueckner, Ed Coulson, George Galster, Joe Gyourko, Matt Kahn, Duke Kao, Janet Kohlhase, Chris Mayer, Dan McMillen, Steve Ross, and William Strange. I thank seminar participants at the University of British Columbia, University of Connecticut, University of Pennsylvania, University of Southern California, University of Wisconsin, Fannie Mae, Federal Reserve Bank of Chicago, Federal Reserve Bank of Kansas City, Homer Hoyt Advanced Studies Institute, and the MacArthur Foundation. Excellent research assistance has been provided by Yong Chen, Ana Dammert, Michael Eriksen, Peter Howe, and Jerry Kalarickal. Any errors are mine alone.

Appendix A. Data creation, summary statistics, and supplemental tables

Three panels were developed to conduct the analysis in this paper, each of which follows neighborhoods over time based on an inter-temporally consistent set of geographic boundaries. The first panel combines census tract data on a decade-by-decade basis from 1950 through 2000 for 35 metropolitan areas in the United States in 1950.[30] Because some of these tracts were subsequently reassigned by Census to other MSAs, this same group of tracts belongs to a broader group of MSAs as of the year 2000. When forming the balanced panel of census tracts from 1950 to 2000 only those tracts belonging to the same MSA in both 1950 and 2000 were retained. This applies to all instances in which the balanced panel is referenced earlier in the paper. In addition, data for Portland Oregon in 1960 were not obtained and for that reason, in some exercises 35

[29] While many of Harlem's residents welcomed the President's move as part of a recent revival of the area, others have voiced concerns about skyrocketing rents that are forcing long-term poor residents from their homes. See the USA Today (July 31, 2001, page 1), "I'm home: Clinton opens Harlem Office," by Charisse Jones.

[30] Beginning with 1950, Census adopted the census tract as the geographic unit used to organize the data. For additional details, see *Procedural History of the 1940 Census of Population and Housing*, by Jenkins (1985).

Table A.1
Data sources

Variables	MSAs	Data source
1900 Census Ward-level traits	Philadelphia	IPUMS Minnesota Population Center, University of Minnesota, http://www.ipums.umn.edu
1920 Census Ward-level traits	Philadelphia	IPUMS Minnesota Population Center, University of Minnesota, http://www.ipums.umn.edu
1950 Census tract-level traits	All 36 MSAs in Table A.1	ICPSR #2931 Elizabeth Mullen Bogue (Bogue, 2000b) file, http://www.icpsr.umich.edu/index.html
1960 Census tract-level traits	All 36 MSAs in Table A.1	ICPSR #2932 Elizabeth Mullen Bogue (Bogue, 2000a) file, http://www.icpsr.umich.edu/index.html
1970 to 1990 Census tract-level traits	All 36 MSAs in Table A.1	Geolytics, 40-years CD, http://www.geolytics.com
1950 to 1960 census tract correspondence	All 36 MSAs in Table A.1	*US Censuses of Population and Housing: 1960 Census Tract Final Reports*. Washington, DC: US Dept. of Commerce, Bureau of the Census
1960 to 1970 census tract correspondence	All 36 MSAs in Table A.1	*US Census of Population and Housing: 1970 Census Tract Final Reports*. Washington, DC: US Dept. of Commerce, Bureau of the Census
1990 to 2000 census tract correspondence	All MSAs	Census tract relationship files on the US Census website: P:\urban_ssrosent\Filtering\correspondences\CensusCorr.-1990–2000\Census 2000 Census Tract Relationship Files.htm
1900 and 1920 Ward maps in Raster format	Philadelphia	http://www.library.upenn.edu/census/images/1899a.jpg
1970 boundary files for census tracts	All 36 MSAs in Table A.1	Geolytics, 40-years CD, http://www.geolytics.com
1990 boundary files for census tracts	All MSAs	US Census Bureau Cartographic Boundary Files: http://www.census.gov/geo/www/cob/
2000 boundary files for census tracts	36 MSAs in Table A.1	Geolytics, 40-years CD, http://www.geolytics.com
Public Housing data for 1940 through 2000	All MSAs	These data were obtained from the Department of Housing and Urban Development (HUD).
LIHTC Housing Data	All MSAs	HUD website: http://www.huduser.org/datasets/lihtc.html

cities are used whereas in other instances only 34 cities are examined.

An additional census tract panel was obtained from Geolytics Inc. (http://www.geolytics.com). This panel follows all identified census tracts in the US on a consistent geographic basis from 1970 to 2000 (their neighborhood change database). For this latter panel, it is important to recognize that census tract coverage of the US expanded over time and covered the entire country for the first time in 1990. These data were further augmented by a third panel for just Philadelphia County from 1900 to 2000. That panel follows each of the 39 year 1900 voting wards in Philadelphia on a consistent geographic basis over time.

Information about the census tract characteristics in each decade were obtained from various sources that are documented in Table A.1 of this appendix. The more challenging data task was to create consistent geographies over time that allow one to follow individual neighborhoods from one decade to the next. This is necessary because census tract geographic boundaries change with each decade. For 1950, 1960, and 1970, correspondence tables from Census data pamphlets are available in printed format for each city in the sample (see Table A.1 for details). Those correspondences indicate the set of tracts from one decade that comprise a tract in the following decade. The Census tables do not, however, report the *degree* to which a tract from an earlier year contributes to one from a later year, only whether tracts from different years overlap. Accordingly, all tracts from a prior decade that contribute to a current-year tract are given equal weight when constructing the correspondence tables that map 1950 tracts to 1960 tract geography, and 1960 tracts to 1970 geography. This is an approximation, but it is the best that can be done for these years.

For the years from 1970 through 2000, Geolytics has coded decennial census tract data for 1970, 1980, 1990, and 2000 in terms of year-2000 tract boundary geography. This was done by matching geography using block-level boundary files available from Census for the last few decades. Because block-level geography is smaller than that of individual tracts, the correspondences for the later decades are more precisely measured than for the earlier years.

Combining the Geolytics data with those that I created for 1950, 1960, and 1970 yields two distinct sets of data: tract data for the years 1970 through 2000 coded to year-2000 Census tracts, and tract data for the years 1950 and 1960 coded to year-1970 Census tract boundaries. It was necessary, therefore, to convert the 1970 tract boundaries to those of year 2000 so that the entire data set can be specified based on the same set of geographic areas. Although Geolytics does not make available the weight matrices to make this conversion, they do make available an electronic map of the 1970 Census tracts. Combining that map with an electronic map of the year-2000 Census tract boundaries from Census it was possible to use mapping software to compute weights that convert 1970 tract boundaries into year-2000 tracts. This was done using MapBasic, the programming language that underlies MapInfo, in conjunction with MapInfo itself to match the geographies for 1970 and 2000. The procedure above was used to produce data based on consistent census tract geography for MSAs across the United States, with all tracts specified in terms of year-2000 tract geography. A second panel was created using all MSAs identified in the Geolytics Census tract data from 1970 to 2000.

The complete list of the 35 metropolitan areas associated with tracts followed from 1950 to 2000 is provided in Table A.2. Also indicated in that table is the number of census tracts identified in each metropolitan area in both 1950 and 2000. Sample means by decade for variables used in the analysis for each of the two panels—the 35 MSA balanced panel and the 1970–2000 "all" MSA panel—are provided in Table A.3.

The final data creation effort focused on Philadelphia County. For this one county, a balanced panel of neighborhood attributes was created for select decades from 1900 to 2000, where neighborhoods for this panel were converted to 1900 Ward-level geographic boundaries. Correspondences that link the early decades with those from more recent years was made possible by very detailed raster image maps that are available for the 1899 and 1914 Ward boundaries in Philadelphia County (see Table A.1 for details). Those maps were converted to electronic boundary files using MapInfo. Correspondences were then developed using MapBasic and MapInfo in order to code data for all of the decades used in the Philadelphia County analysis from Section 2 to year 1900 Ward boundary geography.

An additional challenge in working with the 1900 and 1920 decades for Philadelphia is that Ward-level income was not directly reported by Census. Instead, as a proxy, I used a constructed variable created by the Integrated Public Use Microdata Series (Ruggles and Sobek, 1997) available at http://www.ipums.org referred to as OCCSCORE. This variable is an estimate of the income that an individual in a given decade would have earned given their occupation if they were otherwise observed in 1950. This variable was used to compute the relative Ward-level income in comparison to the Philadelphia County average income for both 1900 and 1920. From 1950 and beyond, income was directly reported in the data and no such approximation was required. Nevertheless, given the approximate nature of the income data from 1900 and 1920, some caution should be used in reviewing the Philadelphia County analysis in Section 2.

Table A.2
Census tracts drawn from 35 1950 MSAs

MSA name	Code	Number of tracts in 1950	Number of tracts in 2000
Akron	80	90	165
Atlanta	520	159	657
Baltimore	720	480	619
Birmingham	1000	92	195
Boston	1120	582	682
Buffalo	1280	191	291
Chicago	1600	1247	1831
Cincinnati	1640	264	399
Cleveland	1680	419	702
Columbus	1840	310	371
Dallas	1920	327	689
Dayton	2000	182	241
Denver	2080	195	508
Detroit	2160	811	1254
Duluth	2240	41	83
Flint	2640	112	130
Hartford	3280	155	287
Houston	3360	651	765
Indianapolis	3480	191	339
Kansas	3760	178	490
Louisville	4520	96	240
Milwaukee	5080	326	411
New Orleans	5560	327	391
Philadelphia	6160	1452	1304
Pittsburgh	6280	361	701
Portland	6440	29	420
Providence	6480	58	257
Richmond	6760	68	252
Rochester	6840	112	263
St Louis	7040	348	519
Seattle	7600	270	522
Syracuse	8160	141	208
Toledo	8400	127	163
Trenton	8480	32	72
Washington	8840	271	1020

A.1. Additional details

It should be noted that for many cities there was a major overhaul of the tract numbering system prior to the 1970 census. Tract names had included letter suffixes prior to this time. For the 1970 census and thereafter tracts were given numerical designations only. Another overhaul of the tract numbering system occurred prior to the 1960 census. All tracts with multiple let-

Table A.3
Sample means

	35 MSA balanced panel 1950–2000						All MSA sample			
	1950	1960	1970	1980	1990	2000	1970	1980	1990	2000
Tract Relative Income (y_t)	1.02	1.02	1.00	0.98	0.97	0.96	1.00	1.00	1.00	1.00
Log(y_t)	−0.01	−0.01	−0.05	−0.09	−0.14	−0.15	−0.04	−0.05	−0.08	−0.09
Log(y_t) − Log(y_{t-10})	–	0.00	−0.04	−0.03	−0.05	−0.01	−0.02	0.00	−0.03	−0.01
% Homes 0–9 yrs	0.28	0.33	0.24	0.17	0.11	0.08	0.33	0.29	0.21	0.15
% Homes 10–19 yrs	0.14	0.14	0.23	0.17	0.14	0.08	0.24	0.20	0.20	0.14
% Homes 20–29 yrs	0.22	0.13	0.13	0.20	0.16	0.13	0.12	0.17	0.16	0.17
% Homes 30–39 yrs	–	–	–	0.14	0.19	0.15	–	0.11	0.15	0.15
% Homes 30 or more yrs	0.36	0.41	0.41	0.46	0.60	0.72	0.31	0.33	0.43	0.54
% Homes 40 or more yrs	–	–	–	0.32	0.41	0.57	–	0.23	0.28	0.39
Density (1000 units/mile)	3.54	3.52	2.81	2.88	2.86	2.85	2.00	2.21	2.35	2.52
Public housing (100s)	0.00	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.01
LIHTC housing (100s)	–	–	–	–	–	0.01	–	–	0.00	0.01
Homeownership Rate	0.64	0.69	0.64	0.62	0.61	0.62	0.67	0.65	0.63	0.64
% Less than high school	0.60	0.55	0.46	0.34	0.27	0.22	0.44	0.31	0.26	0.21
% High school or some college	0.32	0.36	0.42	0.49	0.50	0.51	0.44	0.51	0.51	0.52
% College degree or more	0.08	0.09	0.12	0.18	0.23	0.27	0.12	0.18	0.23	0.27
% Adults married	0.71	0.72	0.64	0.57	0.50	0.46	0.66	0.61	0.55	0.51
% Age 15 or under	0.26	0.31	0.28	0.21	0.21	0.21	0.29	0.22	0.21	0.21
% Age 16 to 29	0.22	0.19	0.24	0.28	0.23	0.21	0.24	0.27	0.23	0.21
% Age 30 to 54	0.36	0.33	0.29	0.29	0.34	0.37	0.29	0.30	0.35	0.37
% Age 55 or over	0.16	0.17	0.19	0.22	0.22	0.21	0.18	0.20	0.21	0.21
% Not African American or Hispanic	0.93	0.90	0.83	0.76	0.71	0.64	0.85	0.81	0.77	0.71
% African American	0.07	0.10	0.15	0.20	0.23	0.27	0.09	0.12	0.14	0.16
% Hispanic	0.00	0.00	0.02	0.04	0.06	0.09	0.06	0.07	0.09	0.13
Income inter-quartile range ($1000)[a]	20.67	12.18	39.28	36.83	43.85	49.05	37.26	35.25	43.04	48.22
Distance to CBD (miles)	9.62	9.62	9.62	9.62	9.62	9.62	11.85	12.25	12.66	12.66
Number of tracts	9419	9419	9419	9419	9419	9419	43,523	48,950	50,312	50,511

[a] Data from each decade are in year 2000 dollars.

ter suffixes used in the 1950 census were renumbered to eliminate them at this time. While no documentation of the letter to number coding was found, in the 1960 data from ICPSR #2932 (see Table A.1) both forms of tract codes were provided. This allowed me to decipher the format used for coding the letters into numbers in the other data sets. Letters "a" through "z" were given values of 01 through 26, "aa" through "zz" were given values of 27 through 52, "aaa" through "zzz" was assigned numbers 53 through 78. No letter suffix's greater than "sss" was found in any of the cities used in the sample.

Further details for each of the steps necessary to create the 1950 to 2000 data sets are provided below.

A.2. Tract correspondences 1950 to 2000

Tract level data for 1950 and 1960 were taken from the Bogue files (see Table A.1). As part of those files, each tract was referenced by an ID. Those codes were then used in conjunction with comparability tables in the MSA-specific Census pamphlets (see Table A.1) to recode the data to a consistent set of geographic boundaries. For 1950, the first two columns of the Bogue files contain the codes that identify the 1950 tracts. These tract ID columns look something like the following:

00130 13
00140 14
00150 15
00171 17 → This number is 17A in the tract comparability charts
00172 18 → This number will be read as 17B in the tract comparability charts.
00173 19

The first column is the column of census tract reference ID numbers, referred to in the 1950 Bogue files as the "tractid." The last number in the tractid is a subscript to the tract ID code. If the value of the last number is 0, then it can be ignored. For instance, 140 is tractid 14. However, if the last number is non-zero, then in the Census MSA pamphlets the non-zero code is represented in alphanumeric code as follows: 171 in the Bogue files appears as 17A in the Census pamphlet, while 172 is 17B, and so on. Note also, that the second column of numbers above are referred to as the "tractseq" in the 1950 Bogue files. These numbers are

Table A.4
Change in neighborhood economic status between 1990 and 2000 with lagged controls.
Dependent variable: $\log(y_{2000}/y_{1990})$

	35 MSAs With Covariates from ...					All MSAs With Covariates from ...		
	1990	1980	1970	1960	1950	1990	1980	1970
% Homes 0 to 9 yrs	0.0490	−0.0577	−0.0990	−0.0415	−0.1076	0.0204	−0.0588	−0.0868
	(3.00)	(3.76)	(6.98)	(2.72)	(6.54)	(3.32)	(8.32)	(13.38)
% Homes 10 to 19 yrs	−0.0595	−0.0892	−0.0256	0.0404	−0.0817	−0.0445	−0.0885	−0.0510
	(3.51)	(5.58)	(1.79)	(2.36)	(3.62)	(6.63)	(11.43)	(7.66)
% Homes 20 to 29 yrs	−0.0634	0.0080	0.0180	−0.0787	−0.0503	−0.0722	−0.0272	0.0055
	(4.14)	(0.53)	(0.82)	(4.92)	(3.36)	(9.84)	(3.41)	(0.51)
% Homes 30 to 39 yrs	−0.0253	0.0091				−0.0566	−0.0067	
	(1.66)	(0.37)				(7.14)	(0.50)	
% Public housing units	−0.1543		0.0025	−0.0004	−0.0023	−0.0080	0.0007	0.0014
	(8.27)		(0.08)	(0.12)	(0.12)	(0.96)	(0.07)	(0.11)
% LIHTC housing units	−0.0595	−0.1341				−0.0507		
	(0.73)	(4.97)				(1.30)		
Density (1000 units/mile)	−0.0065	−0.0056	−0.0034	−0.0019	−0.0019	−0.0030	−0.0045	−0.0031
	(8.54)	(6.96)	(4.12)	(3.22)	(4.23)	(11.95)	(13.21)	(8.72)
Homeownership rate	0.1867	0.0254	−0.0183	−0.0448	−0.0329	0.2302	0.0640	0.0157
	(12.40)	(1.58)	(1.16)	(2.76)	(2.23)	(34.79)	(8.51)	(2.17)
% High school degree (heads)	0.0408	−0.0050	0.0544	−0.0450	0.0486	−0.0212	−0.0467	−0.0089
	(1.27)	(0.16)	(1.81)	(1.41)	(1.49)	(1.54)	(3.18)	(0.65)
% College degree+ (heads)	0.6514	0.4025	0.3265	0.3365	0.3964	0.4916	0.2557	0.2603
	(22.35)	(13.73)	(9.64)	(9.12)	(8.06)	(38.05)	(18.03)	(16.56)
% Married (men 18+)	−0.2422	−0.3277	−0.3132	−0.2816	−0.0936	−0.1698	−0.2528	−0.1963
	(7.23)	(8.94)	(9.78)	(6.43)	(2.06)	(11.70)	(14.50)	(13.33)
% Pop < age 15	−0.2800	0.1553	0.1637	0.1828	0.0891	−0.4372	0.0956	0.1528
	(4.01)	(1.81)	(2.00)	(2.25)	(1.10)	(13.91)	(2.38)	(4.15)
% Pop age 15 to 29	−0.6231	−0.4803	−0.2448	−0.2104	−0.1171	−0.6271	−0.3414	−0.2022
	(12.88)	(8.07)	(3.76)	(2.85)	(1.42)	(31.56)	(13.06)	(7.32)
% Pop age 55+	−0.2653	−0.0508	0.0490	0.0658	−0.0320	−0.3890	−0.0653	0.0396
	(5.30)	(0.81)	(0.67)	(0.78)	(0.31)	(19.53)	(2.47)	(1.36)
% African American (heads)	−0.1056	−0.0649	−0.0378	0.0024	0.0289	−0.1106	−0.0566	−0.0242
	(11.14)	(6.18)	(3.60)	(0.22)	(2.03)	(20.73)	(9.53)	(4.08)
% Hispanic (heads)	0.0047	0.0807	0.1201	0.1595	1.8355	−0.0387	−0.0050	0.0043
	(0.21)	(2.87)	(2.78)	(2.04)	(3.60)	(4.22)	(0.47)	(0.36)
IQR Log Inc (year 2000 $)	0.0017	−0.0002	0.0000	0.0010	−0.0007	0.0018	0.0000	0.0000
	(9.76)	(0.98)	(0.22)	(1.23)	(1.83)	(22.61)	(0.05)	(0.31)
Distance to 2000 CBD (miles)	0.0025	0.0023	0.0024	0.0018	0.0013	0.0013	0.0012	0.0011
	(7.12)	(6.38)	(7.00)	(5.34)	(3.86)	(13.78)	(11.39)	(11.20)
$\log(Y_{1980})$		−0.2082	−0.1360	−0.1055	−0.0988		−0.1732	−0.1290
		(10.90)	(7.34)	(9.57)	(11.10)		(20.21)	(15.50)
$\log(Y_{1980}) - \log(Y_{1970})$		0.1037	0.1097	0.0844	0.0719		0.0791	0.1132
		(6.92)	(5.61)	(5.36)	(4.83)		(12.24)	(13.02)
$\log(Y_{1990})$	−0.4183					−0.3993		
	(26.82)					(60.70)		
$\log(Y_{1990}) - \log(Y_{1980})$	−0.0719					−0.0802		
	(4.94)					(13.67)		
Observations	9419	9419	9402	9406	9419	48,622	43,654	43,397
MSA fixed effects	34	34	34	34	34	325	27	270
R-squared (within)	0.19	0.09	0.07	0.06	0.06	0.16	0.06	0.05

Absolute value of t-ratios in parentheses based on robust standard errors.

also used to identify the individual tracts in the Bogue data for 1950 but they are not codes provided by Census per se. In the 1960 Bogue files census tracts are coded in a manner very similar to the structure of the 1950 data.

Census Bureau pamphlets (see Table A.1) were then used to identify the "tractid" correspondences for those tracts whose boundaries changed during the decade. Tracts not listed in these tables did not change boundaries during the decade and the 1960 tract remains the same as the comparably named 1950 tract. These data were then used to convert all of the 1950 tract attributes into 1960 tract-level geography.

Converting data from 1960 tract geography to 1970 tract boundaries was much easier. In part, this is because comparability charts from the MSA-specific census tract pamphlets for 1970 (see Table A.1) provide the complete set of correspondences between the tracts for 1970 and 1960.

Data for 1970, 1980, 1990, and 2000 were obtained from the Neighborhood Change Database Census CD marketed by Geolytics Inc. (http://www.geolytics.com). As discussed above, these data were all precoded by Geolytics to year-2000 tract boundaries. In addition, as noted earlier, Geolytics did not provide the correspondence weights necessary to convert year 1970 data to year 2000 tract geography. However, Geolytics did provide electronic boundary files for 1970 tracts. These boundary files were matched to the year 2000 tract boundary files to create the correspondences using mapping software (MapBasic and MapInfo). This produced the correspondences that convert 1970 tracts to year 2000.

A.3. Ward correspondences for 1900 and 1920

For Philadelphia County, I started with the Ward districts for 1900 and 1920. Raster image maps showing the ward boundaries for these years were available over the web (see Table A.1). These maps were imported into MapInfo and oriented using county borders as references by overlaying a boundary file for counties in the United States. The county boundary file was obtained from MapInfo (and ultimately, Census). Ward boundaries were then converted to data using drawing tools in MapInfo. Correspondence to 2000 tracts were then computed using mapping techniques in conjunction with year 2000 tract boundary files obtained from the Census website (see Table A.1).

A.4. Public housing and low income hosing tax credit (LIHTC) housing

Data on public housing units in 1990 were obtained from the Department of Housing and Urban Development (HUD). Data for Low Income Housing Tax Credit (LIHTC) housing units were downloaded from HUD's website (see Table A.1). Because LIHTC program commenced in 1986, the number of such units is zero in all decades prior to 1980. For that reason, this variable is used only in those few specifications for which 1990 covariates are entered into the model.

References

Aaronson, D., 2001. Neighborhood dynamics. Journal of Urban Economics 49, 1–32.

Arnott, R., Braid, R., 1997. A filtering model with steady-state housing. Regional Science and Urban Economics 27, 515–546.

Baer, W., 1986. The shadow market in housing. Scientific American 255, 29–35.

Baltagi, B.H., 2005. Econometric Analysis of Panel Data. Wiley, West Sussex, England.

Bogue, D., 2000a. Census tract data, 1960: Elizabeth Mullen Bogue file [Computer file]. ICPSR version. University of Chicago, Community and Family Study Center [producer], 1975. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Bogue, D., 2000b. Census tract data, 1950: Elizabeth Mullen Bogue file [Computer file]. ICPSR version. University of Chicago, Community and Family Study Center [producer], 1975. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Bond, E., Coulson, E., 1989. Externalities and neighborhood change. Journal of Urban Economics 26, 231–249.

Brueckner, J., 1977. The determinants of residential succession. Journal of Urban Economics 4, 45–59.

Brueckner, J., 1980. Residential succession and land-use dynamics in a vintage model of urban housing. Regional Science and Urban Economics 10, 225–240.

Brueckner, J., Rosenthal, S., 2006. Gentrification and neighborhood housing cycles: Will America's future downtowns be rich? Working paper.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge Univ. Press, Cambridge.

Coulson, E., Bond, E., 1990. A hedonic approach to residential succession. Review of Economics and Statistics, 433–443.

Cragg, J.G., Donald, S.G., 1993. Testing identifiability and specification in instrumental variables models. Econometric Theory 9, 222–240.

Cummings, J.L., DiPasquale, D., 1999. The low-income housing tax credit: The first ten years. Housing Policy Debate 10, 257–267.

Cummings, J.L., DiPasquale, D., Kahn, M., 2002. Measuring the consequences of promoting inner city homeownership. Journal of Housing Economics 11, 330–359.

DiPasquale, D., Glaeser, E., 1999. Incentives and social capital: Are homeowners better citizens? Journal of Urban Economics 45, 354–384.

Dye, R.F., McMillen, D.P., 2007. Teardowns and land values in the Chicago metropolitan area. Journal of Urban Economics 61, 45–63.

Eriksen, M., Rosenthal, S.S., 2007. Crowd out, stigma, and the effect of place-based subsdized rental housing. Working paper. Syracuse University.

Galster, G., 1996. William Grigsby and the analysis of housing submarkets and filtering. Urban Studies 33, 1799–1805.

Geolytics website: www.geolytics.com.

Glaeser, E.L., Gyourko, J., 2005. Urban decline and durable housing. Journal of Political Economy 113, 345–375.

Glaeser, E.L., Sacerdote, B., 1999. Why is there more crime in cities? Journal of Political Economy 107, 225–258.

Green, R.K., White, M., 1997. Measuring the benefits of homeowning: Effects on children. Journal of Urban Economics 41, 441–461.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Harding, J., Sirmans, C.F., Rosenthal, S.S., 2007. Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. Journal of Urban Economics 61, 193–217.

Harris, R.D., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91, 201–226.

Haurin, D.R., Parcel, T., Haurin, J., 2002. Does homeownership affect child outcomes? Real Estate Economics 30, 635–666.

Jenkins, R.M., 1985. Procedural History of the 1940 Census of Population and Housing. Madison Univ. of Wisconsin Press.

Levin, A., Lin, C.F., Chu, C., 2002. Unit root test in panel data: Asymptotic and finite sample properties. Journal of Econometrics 108, 1–25.

Maddala, G.S., Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple test. Oxford Bulletin and Economics and Statistics 61, 631–652.

Murray, M.P., 2006. Avoiding invalid instruments and coping with weak instruments. Journal of Economic Perspectives 20, 111–132.

Office of Housing and Community Development (OHCD), City of Philadelphia, 1997, Year 23 Consolidated Plan (Fiscal Year 1998). City of Philadelphia Report: 9.

Ohls, J.C., 1975. Public policy toward low income housing and filtering in housing markets. Journal of Urban Economics 2, 144–171.

Olsen, E.O., 2003. Housing programs for low-income households. In: Moffitt, R.A. (Ed.), Means-Tested Transfer Programs in the US. Univ. of Chicago Press, Chicago. National Bureau of Economic Research.

Pesaran, H., 2003. A simple panel unit root test in the presence of cross section dependence. Cambridge Working Papers in Economics 0346. Faculty of Economics (DAE), University of Cambridge.

Rohe, W.G., McCarthy, G., Van Zandt, S., 2000. The social benefits and costs of homeownership: A critical assessment of the research. Working paper No. 00-01. Research Institute for Housing America.

Rosenthal, S., 1988. A residence time model of housing markets. Journal of Public Economics 36, 87–109.

Rosenthal, S., Helsley, R., 1994. Redevelopment and the urban land price gradient. Journal of Urban Economics 35, 182–200.

Rothenberg, J., Galster, G.C., Butler, R.V., Pitkin, J., 1991. The Maze of Urban Housing Markets: Theory and Evidence. The Univ. of Chicago Press, Chicago, IL.

Ruggles, S., Sobek, M., 1997. Integrated Public Use Microdata Series: Version 2.0. Historical Census Projects, Minneapolis, University of Minnesota.

Schelling, T., 1971. Dynamic models of segregation. Journal of Mathematical Sociology 1, 143–186.

Smith, B.C., 2004. Economic depreciation of residential real estate: Microlevel space and time analysis. Real Estate Economics 32, 161–180.

Staiger, D., Stock, J., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.

Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D., Stock, J. (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge Univ. Press, Cambridge.

Sweeney, J.L., 1974. A commodity hierarchy model of the rental housing market. Journal of Urban Economics 1, 288–323.

Weicher, J., Thibodeau, T., 1988. Filtering and housing markets: An empirical analysis. Journal of Urban Economics 23, 21–40.