Do “Capitalization Effects” for Public Goods
Reveal the Public's Willingness to Pay?

By NICOLAI V. KUMINOFF AND JAREN C. POPE*

This paper develops a welfare theoretic framework for interpreting evidence on the impacts of public programs on housing markets. We extend Rosen’s hedonic model to explain how housing prices capitalize exogenous shocks to local public goods and externalities. The model predicts that trading between heterogeneous buyers and sellers will drive a wedge between these “capitalization effects” and welfare changes. We test this hypothesis in the context of changes in measures of school quality in five metropolitan areas. Results from boundary discontinuity designs suggest that capitalization effects understate parents' willingness to pay for public school improvements by as much as 75%.

* Kuminoff: Arizona State University, Dept. of Economics, Tempe, AZ 85287 (e-mail: kuminoff@asu.edu). Pope: Brigham Young University, Dept. of Economics, Provo, UT 84602 (e-mail: jaren_pope@byu.edu). We appreciate helpful comments from Pat Bayer, Kelly Bishop, Richard Carson, Janet Currie, Tom Downes, Roger von Haefen, Matt Kahn, Mike Keane, Michael Hanemann, Glenn MacDonald, Gilbert Metcalf, Alvin Murphy, Derek Neal, Ray Palmquist, Ed Prescott, Jonah Rockoff, Ed Schlee, Chris Timmins, Jeff Zabel, and especially V. Kerry Smith. We also thank participants of the 2010 AEA meetings, the 2010 W2133 meetings, and the 2010 Duke conference on housing market dynamics, and seminar participants at Arizona State University, Brigham Young University, Research Triangle Institute, Tufts University, University of Calgary, University of Tennessee, Utah State University, Virginia Tech, and Washington University in St. Louis.
1. Introduction

In his seminal 1974 paper, Sherwin Rosen explained how market transactions can reveal buyers’ willingness to pay for the characteristics of a differentiated product. Rosen’s model is frequently used to assess the benefits of policies targeting public goods and externalities. The logic is simple. Homebuyers implicitly purchase the right to consume a bundle of local public goods when they buy a house in a certain neighborhood. It follows that a hedonic price function for housing can be used to infer buyers’ willingness to pay for policies that would alter the provision of public goods. Unfortunately, this is easier said than done.

While there are several complications with using housing market outcomes to estimate the willingness to pay for public goods, recent studies have highlighted endogeneity problems that arise from the market clearing process. As heterogeneous households sort themselves across an urban area they also vote on the provision of local public goods, they interact with their neighbors, and their collective actions may degrade the natural environment (Nechyba 2000, Calabrese et al. 2006, Bayer and Timmins 2007, Walsh 2007, Banzhaf and Walsh 2008, Kuminoff, Smith, and Timmins 2010). This sorting process may produce an equilibrium in which several different public goods are spatially correlated. When data are not available for one or more of these public goods, conventional hedonic estimators are thought to suffer from omitted variable bias.

A recent wave of empirical research has sought to mitigate the omitted variable problem by placing hedonic price functions within the econometric framework for program evalua-

---

1 Other issues include the form of the equilibrium price function and the challenges with estimating demand curves for product characteristics (see Epple 1987, Cropper, Deck, and McConnell 1988, Ekeland, Heckman, and Nesheim 2004, and Kuminoff, Parmeter, and Pope 2010).
tion (Imbens and Wooldridge 2009). The most common strategy is to use a plausibly exogenous source of temporal variation in the quality of a public good to identify how the quality change was capitalized into housing prices. These “capitalization effects” are then interpreted as welfare measures. Researchers have relied on this logic in order to draw strong conclusions about important problems such as the value of a statistical life (Davis 2004), the benefits of the Clean Air Act (Chay and Greenstone 2005), and homeowners’ willingness to pay to reduce their exposure to crime risk (Linden and Rockoff 2008, Pope 2008). More generally, over the past decade the hedonic program evaluation framework has become a leading approach to measuring the public’s willingness to pay for public goods, with numerous applications published in the top general interest and field journals in economics (see Parmeter and Pope 2012 for a survey).

While it is routinely asserted that “capitalization effects” measure the public’s willingness to pay for public goods, no specific evidence has been provided to support this claim. Rosen’s (1974) hedonic model cannot help in this regard. Rosen characterizes a static equilibrium; he does not attempt to explain how markets adjust following exogenous shocks to product attributes. Moreover, none of the studies that have interpreted capitalization effects as welfare measures have developed structural models of the capitalization process to support their interpretations. As Chetty (2009), Heckman (2010), and Keane (2010) all observe, a structural model is usually needed to provide an economic interpretation of a treatment effect.

The purpose of this paper is to investigate the validity of interpreting capitalization effects as welfare measures when the price functions that clear a market for a differentiated
good arise from the equilibrium sorting process described by Rosen (1974). In the first half of the paper we extend Rosen’s conceptual model to express the capitalization effect for a public good as a general function of structural parameters describing household preferences, production technology, and market institutions. We find that the capitalization effect does not have a specific welfare interpretation in this environment. When there is an exogenous shock to the spatial distribution of a public good, the gradient of the hedonic price function will generally adjust in order to clear the housing market. This adjustment drives a wedge between the average capitalization effect and the average household’s willingness to pay.

For example, a reduction in violent crime will change the shadow price of exposure to a given crime rate. The shadow price adjusts because the demand for safety is downward sloping and/or because the composition of households in a given neighborhood changes. A reduction in crime may also change what people are willing to pay for complementary housing attributes such as a home security system or a location near a city park. The problem is that the capitalization effect conflates the public’s willingness to pay for the reduction in crime with changes in the shadow prices of crime and other housing attributes. This type of conflating appears to be a general feature of the hedonic equilibrium model. It even occurs in simple specifications for consumer preferences such as the linear-quadratic-normal model considered by Epple (1987) and Ekeland, Heckman, and Nesheim (2004).

In the second half of the paper we investigate the empirical implications of “conflation bias” in the willingness to pay for public goods. We develop and demonstrate a methodology for testing whether capitalization effects reveal welfare measures. Given a parametric
specification for the hedonic price function, we derive sufficient conditions for interpreting
the marginal capitalization effect experienced by a household as a measure of that house-
hold’s marginal willingness to pay. Importantly, these conditions can be tested within the
hedonic program evaluation framework. Our main test relies on having a research design
for identifying the gradient of the equilibrium price function both before and after the
shock to public goods that defines the capitalization effect.

Our empirical demonstration of the methodology uses a “boundary discontinuity” de-
sign to estimate parents’ valuation of public school quality before and after there were
large changes in publicly reported measures of academic performance. This research
design exploits a series of laws that create spatial discontinuities in the way that children
are assigned to public schools. Children living in physically similar houses in the same
neighborhood are sometimes assigned to different schools where students tend to score
better or worse on standardized exams. These assignment laws underlie our strategy for
estimating the shadow price of school-level academic performance. We estimate shadow
prices in 10 housing markets: five metropolitan areas (Los Angeles, Philadelphia, Detroit,
Fairfax, and Portland) each observed at two points in time (2003 and 2007) that were
chosen because they bracket substantial changes in the measures of test scores that were
reported to parents and the general public. Prior studies such as Black (1999) and Bayer,
Ferreira, and McMillan (2007) have used the same research design to estimate the shadow
price of public school test scores in a single metro area at a single point in time. Their
results provide a baseline for comparison. Our study is the first to provide evidence on
variation in the shadow price of public school test scores across time and space.
We find that the average shadow price of a 1% increase in test scores increased by 28% between 2003 and 2007. This average reflects considerable heterogeneity across metro areas. Changes in the shadow prices of test scores and other housing attributes are conflated with parents’ valuation of public school quality, causing our estimates for capitalization effects to understate hedonic measures of the willingness to pay by as much as 75%.

Overall, the collective evidence from our conceptual and empirical models suggests that the bias in interpreting capitalization effects as measures of the willingness to pay for public goods is of first-order importance. Our work raises the bar for future research. In order to use capitalization effects to draw credible inferences about consumer welfare, the analyst must first demonstrate that the evolution of the hedonic price function supports their interpretation.

The next section provides context for our study and explains our research design. Sections 3 and 4 develop our conceptual and econometric models. Section 5 describes the application to school quality, section 6 presents results, and section 7 concludes.

2. The Hedonic Method and Benefit Measurement

2.1. Identifying Capitalization Effects for Endogenous Public Goods

To illustrate the issues at stake, we begin with a reduced-form model of the relationship between housing prices and public goods. We define “public goods” broadly to include any nonmarket goods and services conveyed to homeowners through their choice of a neighborhood. Examples include local public goods (such as school quality), urban and environmental services (such as crime rates and air quality), and variables describing the demographic composition of the community (such as race and educational attainment).
$P, g, \text{ and } h$ will be used to represent monotonic functions of the price of housing ($P$), the public good of interest ($g$), and all other observable attributes of houses and neighborhoods ($h$). These functions are typically equalities or natural logs (e.g. $P = \ln \text{ price}$). Virtually all reduced-form studies in the hedonic literature pose the following model for the empirical relationship between housing prices and attributes,

$$P_i = g_i \theta_i + h_i \eta_i + \epsilon_i,$$

where subscripts denote the time period, and $\epsilon_i$ is an error term that arises, in part, due to unobserved attributes of neighborhoods (i.e. omitted variables). $\theta_i$ is the parameter of interest. We will discuss its interpretation in section B.

Many public goods are endogenously determined through the housing market in ways that are likely to induce correlation between $g_i, h_i, \text{ and } \epsilon_i$, creating a problem for OLS estimation of (1). For example, imagine trying to isolate the impact of registered sex offenders on nearby property values. If the sex offenders sorted themselves into subdivisions with higher preexisting crime rates, where housing was cheaper, then the OLS estimator for $\theta_i$ will confound the sex offenders’ impact on property values with the impact of preexisting crime. The two effects cannot be distinguished by controlling for crime because data on crime rates are generally unavailable below the level of a zip code. This type of confounding is widely believed to pervade the literature.

Recent studies have developed research designs that mitigate confounding (e.g. Black 1999, Davis 2004, Chay and Greenstone 2005, Linden and Rockoff 2008, Pope 2008, Greenstone and Gallagher 2008, Baum-Snow and Marion 2009, Cellini, Ferreira, and
Rothstein 2010). With the exception of Black (1999), these studies develop identification strategies that exploit a source of temporal variation in $g$. This variation is used to estimate an econometric model specified in terms of fixed effects, first differences, or difference-in-differences. For example, suppose $P$, $g$, and $h$ are observed again after the distribution of $g$ has changed. $h$ and $\varepsilon$ may have changed as well. Differencing the data produces a panel model,

$$
\Delta P = \Delta g \phi + \Delta h \gamma + \Delta \varepsilon ,
$$

where $\Delta P = P_2 - P_1$, for example.$^2$ Equation (2) describes how prices adjusted to the change in $g$, controlling for concomitant changes in $h$. We refer to $\phi$ as the “capitalization effect” because it describes how the change in $g$ was capitalized into housing prices. Many studies use instruments for $\Delta g$ to identify $\phi$. Our interest lies with how these capitalization effects are interpreted.

### 2.2. The Welfare Interpretation of Capitalization Effects

The interpretation of $\phi$ begins with the interpretation of $\theta_i$. It is standard to translate $\theta_i$ into a welfare measure by appealing to hedonic theory. First, equation (1) is assumed to be an equilibrium price function. Differentiating with respect to $g$ defines the marginal price function for $g$. The marginal price paid by the buyer of house $j$ is defined by a $(\theta_i, p_{ij}, g_{ij})$ triplet.$^3$ Next, buyers and sellers are assumed to satisfy the “smoothness conditions” of Rosen’s (1974) model, including: (i) free mobility; (ii) continuity in product

---

$^2$ All of the conclusions in this paper continue to hold if we instead consider the fixed-effects or difference-in-difference analogs to equation (2).

$^3$ The formula for the marginal price depends on whether $P$ and $g$ are measured in levels or logs.
attributes; (iii) perfect information; and (iv) perfect competition. Under these conditions, Rosen demonstrates that the marginal price function for \( g \), evaluated at \( p_{1j}, g_{1j} \), will equal the buyer’s willingness to pay for a marginal change in \( g \) (henceforth MWTP).

In contrast, Rosen (1974) does not interpret \( \phi \). He considers market equilibrium, not the market adjustment process that would follow an exogenous change in product attributes. Studies that estimate capitalization effects have addressed this knowledge gap by assuming that the gradient of the hedonic price function is constant over the duration of the study period (i.e. \( \theta_1 = \theta_2 \) and \( \eta_1 = \eta_2 \) ). This assumption is crucial. It allows household-specific measures of MWTP to be defined by \((\phi, p_{1j}, g_{1j})\) triplets in period 1 and by \((\phi, p_{2j}, g_{2j})\) triplets in period 2. These definitions follow from the interpretation of \( \theta_1 \) and some simple algebra.\(^4\) If we instead consider the “fixed-effects” or “difference-in-difference” analogs to the first-differenced model in (2), we reach the same conclusion: The assumption of a time-constant gradient is crucial to the analyst’s ability to translate the identified parameters of the econometric model into welfare measures.

Thus, under the time-constant gradient assumption (henceforth TCGA) credible estimates for capitalization effects can be translated into credible welfare measures. Recent studies have maintained TCGA to translate capitalization effects into welfare measures for changes in cancer risk (Davis 2004), crime risk (Linden and Rockoff 2008, Pope 2008), hazardous waste (Greenstone and Gallagher 2008), invasive species (Horsch and Lewis 2009), investment in education (Cellini, Ferreira, and Rothstein 2010), low income housing

\(^4\) Write the period 2 price function as \( p_2 = g_2 \theta_2 + h_2 \eta_2 + \epsilon_2 \). Subtracting the period 1 price function from the period 2 price function reduces to the capitalization model in (2) as long as \( \theta_1 = \theta_2 \) and \( \eta_1 = \eta_2 \). Thus, \( \phi = \theta_1 = \theta_2 \).
credits (Baum-Snow and Marion 2009), open space (Bin, Landry, and Meyer 2009), and particulate matter (Chay and Greenstone 2005) to list only a few. In these studies, the gradient is assumed to be fixed for 10 to 20 years, spanning large changes in $g$, $h$, and potentially $\varepsilon$.

Given the importance of developing credible estimates of MWTP for public goods, it is surprising how little is known about the evolution of hedonic price functions. None of the studies invoking TCGA have tested it or provided evidence to validate it. Nor can we find any prior studies that explain what (if anything) must by assumed about preferences in order to guarantee that a hedonic gradient will be invariant to the types of changes in public goods, wealth, and information that occur over 10-20 year periods.

### 2.3. Related Evidence from Previous Studies

Three sets of studies have considered issues that relate to our research question. First, theory papers by Lind (1973) and Starrett (1981) ask whether a policy that alters the distribution of a public good will produce changes in land values that reveal the social benefits of the policy. Their answer is ‘no’, not if heterogeneous households react to the policy by moving. Sieg et al. (2004) reach the same conclusion in a numerical simulation. One might expect their common finding—that price changes do not reveal welfare effects—to extend to our hedonic setting. However, this is purely an intuitive leap. The models developed by Lind, Starrett, and Sieg et al. relax the “smoothness conditions” that support

---

5 For example, Chay and Greenstone (2005, p.418) conclude that their hedonic analysis of air quality “demonstrates that quasi-experimental approaches can be effective in estimating parameters derived from economic models (e.g. MWTP)” and that welfare calculations based on their estimates for capitalization effects, “suggest that the mid-1970s TSPs nonattainment designation provided a $45 billion aggregate gain to homeowners in nonattainment counties.”

6 The length of the study period and the sizes of the changes in variable are dictated by the instruments needed to support the analyst’s preferred identification strategy.
equilibria with a one-to-one mapping between marginal prices and MWTP. Therefore, their results do not have direct implications for the relationship between capitalization and MWTP in environments based on Rosen (1974).

Second, Palmquist (1988, 1992) considers how hedonic price functions could be used to measure welfare effects for changes in environmental quality. His 1988 paper explains how Hicksian welfare measures could, in principle, be constructed from data on an individual’s choices before and after a quality change, if such data were available and if it were possible to identify single-period price functions before and after the quality change. In the special case where the change is “localized”, Palmquist (1992) conjectures that it might be possible to construct welfare measures from knowledge of the ex ante price function. Neither paper addresses the assumptions needed to support TCGA; nor do they consider whether it is possible to recover MWTP from data on the price changes that follow a non-marginal change in quality.

Finally, a few empirical studies have reported evidence of temporal instability in the parameters used to characterize gradients of housing price functions. For example, Brookshire et al. (1985) found that a shock to information about earthquake risk changed the implicit price of earthquake risk over a 6-year period, and Beron, Murdoch, and Thayer (2001) reported annual changes in the implicit price of visibility in Los Angeles between 1980 and 1995. However, the evidence from these studies looks dubious when viewed through the lens of the hedonic program evaluation literature. The problem is that their

---

7 By “localized”, Palmquist means that the quality change has no impact on the equilibrium price function. A single sex offender moving into a neighborhood might fit this definition. On the other hand, subsequent to Palmquist’s work, it has been recognized that localized changes can trigger tipping effects that produce large changes in equilibria (Sethi and Somathathan 2004, Card, Mas, and Rothstein 2008, Banzhaf and Walsh 2010). If the presence of a sex offender induces wealthier white families to move out of the neighborhood, nudging its minority share beyond a tipping point, then there may be large subsequent changes in racial composition and equilibrium prices.
research designs are vulnerable to the omitted variable problems emphasized by Black
(1999), Chay and Greenstone (2005), and others. Thus, previous reports of temporal insta-

bility in hedonic price functions may simply reflect the influence of omitted variables.

In summary: while several authors have considered issues that relate to our research
question, previous studies have not explained the theoretical relationship between capitali-

zation effects and MWTP. Nor have they tested TCGA using a research design that ad-
dresses omitted variable bias.

2.4. Our Research Design

In the next section we extend Rosen’s (1974) conceptual model to derive a relationship
between capitalization effects and MWTP. Our subsequent econometric model character-
izes a bias that arises from using capitalization effects to approximate MWTP if the price
function adjusts to changes in public goods, as the mechanism of Rosen’s model would
suggest. This begs the question: Does “conflation bias” matter for empirical research?
Arguably, the best way to answer this question would be to take a published program
evaluation study, replicate its estimates for capitalization effects, use the ex ante and ex
post cross-sections to test whether the price function changed over time and, finally, quan-
tify any conflation bias. Unfortunately this task is impossible. All of the recent studies
that estimated capitalization effects developed their panel-data identification strategies
because they believed that there was no source of exogenous variation in the public good
capable of identifying a cross-section model. Thus, there are no published studies that
would enable us to simultaneously satisfy best practices for dealing with potential con-
founding in both cross-section and panel-data settings.
Faced with this constraint, we take a second-best approach to evaluating conflation bias. The most direct way to test the hypothesis of a time-constant gradient is to identify single-period price functions before and after a change in the distribution of public goods. While there is no methodological panacea for overcoming confounding in cross-section data, most researchers working in the program evaluation literature appear to view Sandra Black’s (1999) boundary discontinuity design as a credible strategy for mitigating the omitted variable problem. For example, Greenstone and Gallagher (2008 p.997) include it among their short list of papers “demonstrating that it is possible to identify research designs that mitigate the confounding that has historically undermined the credibility of conventional hedonic approaches to valuing nonmarket goods.”

Black (1999) used spatial discontinuities in the laws assigning children to public schools to identify the impact of standardized test scores on property values in the Boston suburbs in the early 1990s. While there have been many subsequent applications of her methodology, none have tracked how the impact of test scores on property values has evolved over time. Nor has any study compared multiple markets at the same point in time. Our study is the first to do both. We estimate 10 hedonic price functions, describing five metropolitan areas, each observed during two years: 2003 and 2007. This period brackets significant changes in the spatial distribution of publicly reported measures of school quality. Details of the data and application begin in section 5. Now we develop the conceptual model.

3. Hedonic Equilibria and the Capitalization of Market Shocks

This section reviews the primitives of Rosen’s model in the context of a housing market,

---

1 Similar opinions are expressed through the discussion of Black’s work in quasi-experimental hedonic studies such as Cellini, Ferreira, and Rothstein (2010), Chay and Greenstone (2005), Linden and Rockoff (2008), and Pope (2008).
characterizes equilibrium, and defines restrictions on preferences and technology that
guarantee the marginal price schedule will be unaffected by exogenous changes in non-
market attributes of a private good. Previous studies have characterized the relationship
between consumer preferences and the hedonic price schedule using closed-form models
of a single equilibrium, absent any exogenous shocks to product attributes. Epple (1987) is
a notable example. He used Tinbergen’s (1959) linear-quadratic-normal model to charac-
terize the identification of demand curves for product attributes in a single equilibrium.9 If
we extend Epple’s analysis to introduce an exogenous shock to a spatially delineated
amenity, it becomes clear that the marginal price schedule will generally need to adjust in
order to clear the housing market. While the linear-quadratic-normal model is elegant, it
does not provide a robust description of market outcomes. Slight modifications to its
structure can produce extremely important changes to the identification of structural pa-
rameters, as Ekeland, Heckman, and Nesheim (2004) demonstrate. Motivated by their
findings, we devote the rest of this section to characterizing the relationship between capi-
talization effects and MWTP in a more generic setting. Readers who would prefer para-
metric and graphical examples are directed to section I of the supplemental appendix,
where we demonstrate our main results in the context of the linear-quadratic-normal mod-
el.

3.1. Demand, Supply, and Market Equilibrium

Price-taking households are assumed to be free to choose a house with any combination
of attributes (e.g. bedrooms, bathrooms, sqft) in the neighborhood that provides their de-

---

9 See Bishop and Timmins (2011) for recent progress on the identification problems described by Epple.
sired levels of public goods (e.g. school quality, air quality). The utility maximization problem is

\[
\max_{g, X, b} U(g, X, b; \alpha) \quad \text{subject to: } y = b + P(g, X; \Theta),
\]

where $X$ denotes all attributes of houses and neighborhoods (other than $g$). A household chooses levels of attributes and the composite good ($b$) to maximize utility, given its preferences ($\alpha$), income ($y$), and the after-tax price of housing, $P(g, X; \Theta)$, which is expressed as a general parametric function of $g, X,$ and a parameter vector, $\Theta$. The first order conditions are

\[
\begin{align*}
(4a) & \quad \frac{\partial P(g, X; \Theta)}{\partial g} = \frac{\partial U/\partial g}{\partial U/\partial b} \equiv D(g; X, \alpha, y), \\
(4b) & \quad \frac{\partial P(g, X; \Theta)}{\partial X} = \frac{\partial U/\partial X}{\partial U/\partial b} \equiv R(X; g, \alpha, y).
\end{align*}
\]

The first equality in (4a) implies that each household will choose a neighborhood that provides a quantity of $g$ at which their willingness to pay for an additional unit equals its marginal implicit price. Assuming the marginal utility of income is constant, the second equality observes that as $g$ varies the marginal rate of substitution defines the inverse demand curve, conditional on $X$. Equation (4b) states analogous first order conditions for $X$.

Let $C(g, M, X; \beta)$ denote a producer’s cost function, where $M$ is the number of houses

---

10 In the context of equation (1), $X$ would include the elements of $h$ along with the omitted variables that enter the error term, $\epsilon_i$. 

they sell and $\beta$ is a vector of cost parameters.\footnote{Producers may include developers, contractors, and individuals selling their homes. For a developer, the cost function will reflect the physical, labor, and regulatory costs of building a home. For a homeowner, the cost function will reflect their psychological attachment to the home as well as the cost of renovation.} Following Rosen (1974), we treat each producer as a price taker who is free to vary the number of units they sell as well as a subset of the characteristics of each unit. For convenience, $g$ is treated as exogenous.\footnote{The results of this section are not altered by making $g$ endogenous or $X$ exogenous. We need only assume that $g$ may be influenced by forces that are exogenous to the model.}

In this case, the profit maximization problem is

\begin{equation}
\max_{X,M} \pi = M \cdot P(g, X; \Theta) - C(g, M, X; \beta),
\end{equation}

with the corresponding first order conditions

\begin{equation}
P(g, X; \Theta) = \frac{\partial C(g, M, X; \beta)}{\partial M}, \quad \frac{\partial P(g, X; \Theta)}{\partial X} = \left( \frac{1}{M} \right) \frac{\partial C(g, M, X; \beta)}{\partial X}.
\end{equation}

Producers choose $M$ to set the offer price of the marginal house equal to its production cost, and they choose $X$ to set the marginal per unit cost of each attribute equal to its implicit price.

Equilibrium occurs when the first order conditions in (4) and (6) are simultaneously satisfied. This system of differential equations implicitly defines the equilibrium hedonic price function that clears the market. It will be useful to rewrite the price function to acknowledge its dependence on model primitives,

\begin{equation}
P(g, X; \Theta) \equiv P[g, X(g, A, B); \Theta(g, A, B)].
\end{equation}

Equilibrium levels of $X$ are determined by all of the exogenous variables: $A : F(y, \alpha) \sim A$, a vector of parameters that describes the joint distribution of household income and prefer-
ences, \( B : V(\beta) \sim B \), a parameter vector describing the distribution of producer characteristics, and \( g \).\(^{13}\)

Importantly, the reduced form parameters describing the shape of the price function are endogenously determined by the structural parameters and exogenous variables, \( \Theta = \Theta(g, A, B) \). Thus, any unexpected changes in the distributions of income, preferences, technology, or public goods can change the shape of the price function which, in turn, changes the implicit price schedule for \( g \).

### 3.2. Interpreting Capitalization Effects as Welfare Measures

Now we depart from Rosen (1974) to consider equilibria in the same geographic market, before and after an exogenous shock to \( g \). The change in the value of house \( j \) depends on the difference in the pre and post-shock price functions,

\[
(8) \quad P[g_{2j}, X_{2j}(g_2, A_2, B_2); \Theta(g_2, A_2, B_2)] - P[g_{1j}, X_{1j}(g_1, A_1, B_1); \Theta(g_1, A_1, B_1)],
\]

where the subscripts denote pre and post-shock equilibria. To isolate the capitalization effect, we condition on \( X \) and divide the change in \( P \) by the change in \( g \),

\[
(9) \quad \phi_j = \frac{P[g_{2j}; \Theta(g_2, A_2, B_2)]X_{2j} = \bar{X} - P[g_{1j}; \Theta(g_1, A_1, B_1)]X_{1j} = \bar{X}}{g_{2j} - g_{1j}}.
\]

This difference quotient provides a general expression for the parameter estimated in the hedonic program evaluation literature.

Because \( \phi_j \) depends on two (potentially different) price functions, it is not the measure

\(^{13}\) \( M \) drops out of the expression for \( X \) in (7) because it is a function of model primitives.
of MWTP from Rosen (1974). To convert $\phi_j$ into MWTP, we must restrict preferences and technology to assure that the capitalization effect will equal the partial derivative of the pre-shock and/or post-shock price functions. Severity of the restriction depends on the size of the shock. If the change in $g$ is small, we need only restrict $A_1 = A_2$ and $B_1 = B_2$. Under this condition, the difference quotient in (9) approaches the partial derivative in (4a) as $g_{2j} - g_{1j}$ approaches zero. In the limit, pre-shock MWTP equals post-shock MWTP which equals the capitalization effect. This is intuitive. An infinitesimal change in a single attribute will not alter the shape of a hedonic price function; equilibrium prices will simply increase by MWTP.

However, as noted earlier, empirical studies typically analyze large shocks. In this case, three restrictions are jointly sufficient to establish a welfare interpretation for the capitalization effect. We state this formally as

**ASSUMPTION 1.**

- $A_1 = A_2$ and $B_1 = B_2$.
- $\partial P(g, X; \Theta)/\partial g = f(X, \Theta)$.
- $\partial \Theta/\partial g = 0$.

Condition $a$ restricts preferences, income, and technology to be constant over the duration of the study. It follows that supply and demand for $g$ and $X$ are also fixed. Condition $b$ implicitly restricts the shapes of supply and demand curves so that the marginal price of $g$ does not depend on its level. Condition $c$ further restricts supply and demand such that changes in $g$ do not affect the hedonic gradient. We discuss violations of each condition.

---

14 Proof of this statement follows immediately from the definition of a derivative.
after proving the theorem.

**THEOREM 1.** If assumption 1 holds for a shock to $g$, then the capitalization effect, $\phi$, reveals the pre-shock MWTP, which equals the post-shock MWTP.

**Proof.** Consider any house, $j$, with characteristics $X_j = \bar{X}$ for which $g_j$ changes from $g_{j1}$ to $g_{j2}$. Since $A_1 = A_2$, $B_1 = B_2$, and $\partial \Theta / \partial g = 0$, we know that $\Theta_1 = \Theta_2$. Combining this result with the assumption that $\partial P(g, X; \Theta) / \partial g = f(X, \Theta)$ implies $f(\bar{X}, \Theta_1) = f(\bar{X}, \Theta_2)$. It follows from the Mean Value Theorem that $\phi = f(\bar{X}, \Theta_1) = f(\bar{X}, \Theta_2)$.

The second term measures pre-shock MWTP and the third term measures post-shock MWTP, as defined by the first-order conditions from Rosen (1974). QED.

Alternatively, if assumption 1 is violated, the Mean Value Theorem generally implies

$$\phi_j \neq \partial P(g_{j1}, X_{j1}; \Theta_1) / \partial g \neq \partial P(g_{j2}, X_{j2}; \Theta_2) / \partial g.$$  

For example, suppose conditions $b$ and $c$ hold, but the shock to $g$ coincides with a shock to income or information, changing what households are willing to pay for $g$. This example violates condition $a$. Since the parameters defining the hedonic gradient depend on preferences and income, $\Theta_2$ may differ from $\Theta_1$, causing the hedonic gradient to adjust, driving a wedge between the capitalization effect and MWTP.

Now suppose conditions $a$ and $c$ hold so that the price function itself is stable (i.e. $A_1 = A_2$, $B_1 = B_2$, and $\Theta_1 = \Theta_2$). Condition $b$ restricts the curvature of its gradient. This restriction avoids problems that can occur if the gradient depends on $g$. To see this, notice that a movement along a nonlinear price function will generally change marginal prices. If an increase from $g_{j1}$ to $g_{j2}$ corresponds to a change in its price, then the capi-
talization effect cannot simultaneously equal ex ante MWTP and ex post MWTP, since the two measures of MWTP differ. The strength of condition $b$ is underscored by Ekeland, Heckman, and Nesheim’s (2004) finding that the hedonic gradient is generically nonlinear in $g$.

Finally, consider condition $c$. The only obvious restriction on market primitives that supports $\partial \Theta / \partial g = 0$ is that the demand for $g$ is perfectly elastic. If the demand is downward sloping, then a positive shock to $g$ will decrease individual MWTP (changing $\Theta$). Utility should also be separable in $g$ and $X$. Otherwise, a shock to $g$ could change the implicit prices of the elements of $X$. If $g$ is the crime rate, for example, we must be willing to assume that changes in crime do not affect the willingness to pay for security systems, fences, or proximity to city parks. These restrictions on own and cross-price elasticities also apply to elements of $X$ that are subject to exogenous shocks. A change in the relative price of any attribute violates $\partial \Theta / \partial g = 0$ and can drive a wedge between MWTP and the capitalization effect for any other attribute.

Conditions $a$, $b$, and $c$ are obviously strong restrictions. They seem unlikely to be satisfied in most applications. If they are violated, then the hedonic gradient may be unstable, producing a wedge between the identified capitalization effect and the policy-relevant measure of MWTP. The magnitude of this “conflation bias” will depend on correlations in the data and the structure of preferences and technology. We provide parametric and

---

15 As noted earlier the three conditions are sufficient, but they are not strictly necessary. It is possible to construct examples where simultaneous violations of two or more conditions are exactly offsetting. This should not diminish their importance. To provide an analogy: selecting the right empirical specification and valid instruments is sufficient, but not necessary, to identify causal parameters in applications of the instrumental variables model. It is possible to construct examples where the bias from invalid instruments is exactly offset by biases from measurement error. This certainly does not diminish the importance of omitted variable bias.
graphical examples in the appendix, using a quadratic specification for utility. However, the common approach to empirical work is to avoid making parametric assumptions about consumer preferences by instead assuming a parametric form for the equilibrium price function. We follow this approach in the next section, deriving an expression for confla-
tion bias in terms of the data and parameters of a standard reduced form model.

4. Sufficient Conditions for Capitalization Based Welfare Measurement

We begin by repartitioning $X$ into observed ($h$) and unobserved ($\varepsilon$) components. Using this partition, the linear price functions that describe market equilibria before and after an unexpected shock to $g$ are $P_1 = g_1 \theta_1 + h_1 \eta_1 + \varepsilon_1$ and $P_2 = g_2 \theta_2 + h_2 \eta_2 + \varepsilon_2$. Recall that $P$, $g$, and $h$ are monotonic functions of the underlying variables; for example, they may be measured in levels or natural logs. Parameter subscripts recognize that the shape of the function may have been altered by the shock to $g$ and by concomitant changes in $h$, $\varepsilon$, $F(y, \alpha)$, and $V(\beta)$. Note that we do not take a stance on approximation error in the use of a linear functional form.\textsuperscript{16} Since virtually all empirical studies use linear models, doing so here allows us to focus attention on the relationship between capitalization effects and MWTP. The results in this section should be viewed as a “best-case” scenario where the price function is specified correctly.

Subtracting the old price function from the new one yields a general time-differenced model,

\begin{equation}
\Delta P = (g_2 \theta_2 - g_1 \theta_1) + (h_2 \eta_2 - h_1 \eta_1) + \Delta \varepsilon .
\end{equation}

In the special case where $\theta_1 = \theta_2$ and $\eta_1 = \eta_2$, equation (10) reduces to the estimator from (2), $\Delta P = \Delta g \phi + \Delta h \gamma + \Delta \varepsilon$.

Applying the Frisch-Waugh Theorem, the relationship between the estimator for the capitalization effect ($\hat{\phi}$) and the price function parameters describing MWTP ($\theta_i, \theta_2$) can be written as:

$$\hat{\phi} = \theta_2 + \frac{r'g}{r'r}(\theta_2 - \theta_1) + \frac{r'h}{r'r}(\eta_2 - \eta_1) + \frac{r'\Delta \varepsilon}{r'r},$$

where $r = \Delta g - \Delta h(\Delta h'\Delta h)^{-1} \Delta h'\Delta g$. Let $z$ denote a valid instrument for $\Delta g$. The IV analog to (11) simply replaces the $\Delta g$'s in $r$ with $\hat{\Delta g} = z(z'z)^{-1}z'\Delta g$.

Equation (11) reports what we can expect to learn about MWTP from estimating (2) when (10) is the true model. The IV estimator for the capitalization effect, $\hat{\phi}_{IV}$, depends on all of the parameters of the price functions that precede and follow the shock, as well as correlations between levels and changes in housing characteristics. The first term to the right of the equality in (11) is a parameter defining MWTP in the new equilibrium. The second term is a “price effect” arising from a change in the implicit price of $g$. The third term is a “substitution effect” arising from changes in the implicit prices of other housing attributes that affect utility and, in some sense, serve as substitutes for $g$. The last term reflects the bias from correlation between changes in observed and unobserved variables.

The implicit price and substitution effects arise when the hedonic price function acts as the market clearing mechanism that Rosen described, adjusting to clear the market follow-

---

17 Applying Rosen’s FOC to the price function defines MWTP for the occupant of house $j$ by a $(\theta_j, p_j, g_j)$ triplet, with the exact formula determined by the monotonic transformations to the variables in the hedonic price function.
ing the change in \( g \) and any changes in \( h, \varepsilon, F(y, \alpha), \) or \( V(\beta) \). Summing the price and substitution effects defines the “conflation bias” in interpreting a capitalization effect as a parameter of a hedonic price function. The direction of the bias is indeterminate. Using \( \hat{\phi}_{IV} \) to calculate MWTP for the occupant of house \( j \) may produce an estimate that falls outside the range of values for the true MWTP for the occupants of \( j \) in the ex ante and ex post equilibria.\(^{18,19} \) To establish a mapping between capitalization effects and welfare measures, some additional restrictions will be needed.

At least two sets of conditions are sufficient to translate capitalization effects into MWTP. The first set of conditions follows directly from assumption 1. If assumption 1 is satisfied, the hedonic gradient must be time-constant. Adding the usual orthogonality restriction on the econometric error term gives us

\[
(12) \text{SUFFICIENT CONDITION 1.} \quad \theta_1 = \theta_2, \quad \eta_1 = \eta_2, \quad \text{and} \quad z, \Delta h \perp \Delta \varepsilon.
\]

Under these restrictions, equation (11) reduces to \( \hat{\phi}_{IV} = \theta_1 = \theta_2 \). In this case the capitalization model (2) can be used to develop an unbiased estimator of ex ante MWTP which equals ex post MWTP. If estimation of single-period price functions is possible, the time-constant gradient assumption can be tested.

The second set of conditions replaces TCGA with additional restrictions on the data. It

\(^{18}\) For example, consider a quality improvement that decreases MWTP without affecting the control variables or their marginal prices: \( \Delta h = \eta_2 - \eta_1 = \Delta \varepsilon = 0 \). In this case, (11) implies that \( \hat{\phi} < \theta_1 < \theta_2 \) if \( \Delta g' g_1 > 0 \). Alternatively, \( \theta_1 < \theta_2 < \hat{\phi} \) if \( \Delta g' g_2 < -\Delta g' g_1 \).

\(^{19}\) From the perspective of welfare measurement, conflation bias is more problematic than the standard econometric complications with interpreting local average treatment effects (LATE). In the presence of heterogeneous treatment effects, LATE can identify parameters that are “structural” in the sense that they are invariant to policy changes operating through \( z \) (see Heckman 2010). In contrast, capitalization effects for public goods are not “policy invariant”. The market clearing function of the hedonic gradient makes capitalization effects endogenous to changes in implicit prices of non-market goods that will, in turn, vary with the policy change operating through \( z \).
can be seen from (11) that $\hat{\theta}_2 = \theta_2$ under the following conditions

(13) **SUFFICIENT CONDITION 2.** $g_1, h, \Delta h \perp z$ and $z, \Delta h \perp \Delta \epsilon$.

If the instrument is randomized in the sense that it is orthogonal to the initial level of the public good, and to the initial levels of the control variables, and to changes in those variables, then the capitalization effect identifies MWTP in the post-shock equilibrium, even if the gradient changes.

The policy relevance of ex post MWTP depends on the nature of the instrument and the evolution of the hedonic gradient. Consider a large improvement in $g$ that drives MWTP to zero. Knowing $\theta_2$ (but not $\theta_1$) does not allow us to distinguish the hypothesis that people are better off from the alternative hypothesis that people were indifferent to the improvement that occurred. Now suppose a second random event causes $g$ to deteriorate, increasing MWTP. Data from periods two and three could be combined to recover $\theta_3$. In principle, $\theta_2$ and $\theta_3$ could be used to develop a linear approximation to the market demand curve for $g$, providing a more credible foundation for policy analysis.\(^{20}\)

In summary, equations (12)-(13) define sufficient conditions for translating capitalization effects into welfare measures.\(^{21}\) While the first-differenced estimator provides a convenient framework for deriving these conditions, they also apply to the fixed-effects and difference-in-difference estimators. In all three frameworks, econometric consistency is established, in part, through an assumption of temporal stability in the hedonic gradient.

\(^{20}\) It is also worth noting that the approach being suggested here would serve as a quasi-experimental capitalization-based analog to the idea that Palmquist’s (1988) originally outlined for using hedonic price functions to calculate welfare measures for quality changes.

\(^{21}\) This is similar to Chetty’s (2009) “sufficient statistics” for quasi-experimental welfare measurement
If this assumption is valid, then the identified parameters can be translated into ex-ante MWTP, which equals ex-post MWTP. If the gradient changes but an instrument randomizes the amenity “treatment”, then a capitalization effect can be translated into ex-post MWTP. However, the scope for using MWTP in policy evaluations depends on the evolution of the gradient. Thus, studying the evolution of price functions is essential to understanding the mapping between capitalization effects and policy-relevant measures of the willingness to pay for public goods.

5. Evidence on the Evolution of Hedonic Price Functions

The 40-year history of research on valuing school quality is a microcosm for the broader literature on using housing markets to value public goods.22 Because a household’s access to a public school has traditionally been determined by whether the household lives in the attendance zone for that school, property values should reflect what parents are willing to pay for their children to attend schools where students score higher on standardized tests. Early studies appeared to confirm this intuition. Then researchers noted a potential source of confounding—schools with higher test scores tend to be located in more exclusive neighborhoods. Subsequent studies refined the research design to mitigate confounding from omitted neighborhood amenities. This work began with Black (1999). She noticed that school quality shifts discretely as one crosses an attendance zone boundary, but other amenities do not (e.g. crime, air quality). Therefore, the composite price effect of all unobserved amenities that are common to houses on both sides of a boundary can be absorbed by a fixed effect for the “boundary zone”. By focusing on sales that occurred near a

---

22 Kain and Quigley (1975) is among the early contributions. Recent applications include Downes and Zabel (2002), Figlio and Lucas (2004), Reback (2005), and Bayer, Ferreira, and McMillan (2007).
boundary and including fixed effects for each boundary zone, Black forced the identification to come from price differentials between similar houses located on opposite sides of a boundary.

Bayer, Ferreira, and McMillan (2007) refined Black’s approach to control for correlation between preferences for schools and preferences for the demographic characteristics of one’s neighbors. The problem stems from sorting. If preferences for school quality are correlated with demographic characteristics, such as race or education, then similar “types” of households will tend to locate in the same attendance zones. This helps to explain why neighborhood racial composition also tends to shift discretely as one crosses an attendance zone boundary. Since prospective homebuyers may care about the characteristics of their neighbors, one must control for the demographic composition of the neighborhood in order to isolate the implicit value of academic performance.23

We use the boundary discontinuity design for valuing school quality to estimate single-year price functions for five metropolitan areas at five-year intervals. Then we calculate MWTP for school quality in each year, test TCGA, and compare estimates for MWTP to capitalization effects following changes in test scores that occurred over the first four years of the No Child Left Behind Act (NCLB). Throughout the application, we follow the data collection and econometric procedures outlined by Black (1999) and Bayer, Ferreiria, and McMillan (2007). Readers are referred to their papers for additional background. The remainder of this section briefly summarizes NCLB and the data sets we have assembled.

---

23 In addition to refining Black’s (1999) reduced-form estimation strategy, Bayer, Ferreira, and McMillan (2007) also develop and estimate a random utility model of sorting behavior.
5.1. *No Child Left Behind*

The No Child Left Behind Act was one of the most sweeping reforms in the recent history of public education. Beginning in 2003, it required states to implement accountability systems that measure student performance in reading and math. Standardized testing is done in grades 3 through 8 and at least once during high school. State test scores are used to determine if each school is making “adequate yearly progress” toward the goal of having every student attain state-specific standards for minimum competency in reading and math by 2014. Schools that do not make adequate yearly progress face a series of repercussions.

While test scores have trended up since NCLB was enacted, its impact on the quality of education has been debated. Advocates argue that school quality will be improved by tracking performance, publicizing results, and sanctioning poorly performing schools. Critics argue that NCLB creates perverse incentives to “teach to the test”, to lower standards, to expel poorly performing students, or even lie when reporting scores. Several authors have investigated these issues. The emerging consensus seems to be that NCLB has improved academic performance, despite flaws in the program. For example, Dee and Jacob (2011) identify the impact of NCLB on test scores from the National Assessment of Education Progress (NAEP). A key feature of their research design is that changes in NAEP scores should be unaffected by the perverse incentives that critics of NCLB have emphasized. They find that NCLB produced large and broad gains in NAEP math scores of 4th and 8th graders, especially in the bottom decile of the achievement distribution.24

These results suggest that the upward trend in NCLB scores is consistent with alternative

---

24 Mean increases in the NAEP math test scores were approximately 1-8 points from the start of NCLB to 2007 for 4th and 8th grade math scores. In a related study, Neal and Schanzenbach (2011) find that NCLB increased reading and math scores for fifth graders in the middle of the achievement distribution in the Chicago Public School system.
metrics for judging school quality.

NCLB scores are the official source of public information about school quality, and they are easily observed. Every school is required to track the share of its students who achieve proficiency in each subject. Results are mailed to parents and posted on websites such as greatschools.org.

5.2. Ten Boundary Discontinuity Designs

We estimate housing price functions for the metropolitan areas of Portland OR, Fairfax County VA, Philadelphia PA, Detroit MI, and Los Angeles CA during the 2003 and 2007 school years. After an exhaustive search over prospective study areas, these five were chosen because they had: (i) a large number of boundary zones; (ii) a large number of housing transactions; and (iii) data on NCLB scores in 2003 and 2007. The five regions also provide geographic diversity.

Black (1999) and Bayer, Ferreira, and McMillan (2007) used elementary school attendance zones as the basis for identification. We use this same approach in Fairfax and Portland, where children are still assigned to schools based on the attendance zones where their parents live. However, school-specific assignment is no longer the norm. Since the mid-1990s, there has been an explosion of state and local regulations mandating “open enrollment” at the school district level. In an open enrollment area, parents are free to send their children to any public school within the school district. There is evidence that parents take advantage of these laws by sending their children to schools outside the elementary attend-

---

25 Boundary discontinuity analysis is extremely data intensive because it discards housing transactions that occur beyond small distances from the school district boundaries.
26 States were not required to start reporting test scores until 2006. Some states did not report scores in 2003.
ance zone where their house is located (Reback 2005). Philadelphia, Detroit, and Los Angeles all have open enrollment policies. For these areas, our identification strategy is based on the relationship between property values and test scores on opposite sides of district boundaries.

Implementing the boundary discontinuity design at the school district level requires taking a weighted average over the scores in each district. This has the advantage of smoothing over idiosyncratic variability in annual school-specific scores. Yet, it also requires extra caution. Property tax rates can vary discretely across districts. District boundaries may also be more likely than attendance zone boundaries to overlap with features of the urban landscape. Therefore, we control for property tax rates and we use visual inspection to exclude boundaries that overlap with discernible landscape features such as rivers and highways.

5.3. Data and Summary Statistics

We assembled data on test scores, neighborhood characteristics, and houses sold during the 2003 and 2007 school years.27 The scores that we use are combined rates of math and reading proficiency reported by states under NCLB. We matched each housing sale with lagged scores for the relevant school or school district.28 Houses sold during the 2003 school year were matched with scores from the 2002 school year, for example. We will refer to the lagged scores as the “2003 score” and “2007 score” from here on.

27 The 2003 school year is defined as October 1, 2003 through September 30, 2004, and the 2007 school year is defined as October 1, 2007 through September 30, 2008. These definitions reflect the fact that NCLB scores and school grades for the preceding year are typically announced at the end of August or the beginning of September. Thus we want to allow time for our proxy for school quality—test scores—to influence home buying decisions.

28 The school quality information was obtained from www.schooldatadirect.org. The combined measure of reading and math is an overall measure (calculated by Standard & Poor’s) that provides an average of the proficiency rates achieved across all reading and math tests, weighted by the number of tests taken for each school or school district.
Table 1 reports the 2003 NCLB scores and 2007-2003 differences for the 10th, 50th, and 90th percentiles of schools in each study area. In Fairfax, for example, math/reading scores in the bottom 10th percentile increased by an average of 11 points (or 14%) with a standard deviation of 8 points. The corresponding changes for the other four areas are all positive and typically large. There are smaller gains (and even losses) at the middle and 90th percentiles. These statistics are consistent with Dee and Jacob’s (2011) finding that NCLB had the biggest impact on schools that began the program with the lowest scores.

The remaining components of the data were collected from various sources. Transaction prices and structural characteristics of every house sold during the 2003 and 2007 school years were purchased from a commercial vendor that assembles the data from public records. Tax rates were calculated using assessment data from public records. Finally, each house was matched with data on the demographic composition of residents living in the Census block group.

Table 2 reports summary statistics for Fairfax County, VA. Columns 1-2 report means and standard deviations for every variable in the final data set. In 2003 the average house sold for approximately $567,000. By 2007 the price had dropped slightly to $563,000. Over this same period, the average test score rose from 83.56 to 84.36. This small change in the average masks considerable heterogeneity across schools (table 1). The average house was 34 years old, with 4 bedrooms, 3 baths, and 2,100 square feet of living

---

29 Scores are not directly comparable across study areas because each state has its own testing system.
30 Annual block group data were obtained from Geolytics. Their data are developed using information from the decennial Census, annual Census surveys, postal records, and actuarial tables of births and deaths.
31 It should be noted that the mean for the 2003 score levels is slightly different than the 2003 score level reported in Table 1 and the corresponding appendix tables. This is because Table 2 scores are weighted by enrollment whereas Table 1 is weighted by housing transactions. In other words, the difference represents the fact that the spatial distribution of housing transactions is not the same as the spatial distribution of enrollments.
area on a 0.4 acre lot. It was located in a block group where 23% of the neighborhood was nonwhite, 24% was under 18 years of age, 85% of houses were owner occupied, 1% of houses were vacant, and 0.37 was the normalized measure of population density. The average ratio of assessed to taxed value called a “tax rate” was 112.

Columns 3-5 summarize the subsample that we use in the boundary discontinuity regressions. Column 3 reports means over houses located within 0.2 miles of a boundary. While this cuts the sample in half, there are almost no changes in the characteristics of the average house (comparing columns 1 and 3). Column 4 reports the difference in mean characteristics of houses located on the “high score” and “low score” sides of a boundary, and column 5 reports T-statistics on the differences. Differences in scores are large and statistically significant whereas differences in housing characteristics tend to be small and insignificant. Like Bayer, Ferreira, and McMillan (2007), we find differences in the racial composition of homeowners on the high and low-score sides of a boundary. This underscores the importance of controlling for demographics during the estimation.

Columns 6-7 report means and standard deviations for the average house in each Census block group. Because there are too few repeated sales of individual houses to estimate a first differenced model using micro data, we use the block group data to estimate capitalization effects for changes in test scores between 2003 and 2007. Notice that aggregation does not substantially change the summary statistics relative to the micro data. Finally, columns 8-9 report correlations between the change in test scores and levels and changes in all other variables.

The Fairfax county data illustrate several features that are common to the data sets for
Portland, Philadelphia, Detroit, and Los Angeles. In particular: (i) variable means are very similar across the full micro, 0.2 mile micro, and block group samples in each metro area; (ii) test scores and racial composition both tend to change discretely across the boundary zones; (iii) changes in test scores are negatively correlated with the baseline level of test scores; and (iv) changes in test scores are generally correlated with levels and changes in other housing characteristics. Summary statistics for each area are provided in the online appendix.

6. Results

6.1. Single-Year Hedonic Regressions

Our hedonic estimates of the MWTP for school quality are based on the following specification for the price function:

\[
\ln P = \text{testscore}\theta_{03} + D \cdot \text{testscore}\theta_{07} + h\eta_{03} + D \cdot h\eta_{07} + BFE_{03} + BFE_{07} + \epsilon.
\]

\text{testscore} denotes the log of the NCLB score for the year prior to the housing sale, \(D\) is an indicator for sales during the 2007 school year, \(h\) includes all structural housing characteristics, neighborhood demographics, and the tax rate, and \(BFE_{03}, BFE_{07}\) are year-specific boundary fixed effects. The boundary regions are 0.2 mile areas that overlap adjacent attendance zones (Fairfax, Portland) or adjacent districts (Philadelphia, Detroit, Los Angeles).\(^{32}\) Under the null hypothesis that the hedonic gradient is constant over the duration of the study, \(\theta_{07} = \eta_{07} = 0\).

We begin by using the sample of houses that sold within 0.2 miles of a boundary. Pan-

\(^{32}\)Our main results are unaffected by using boundary regions of 0.35 or 0.15 miles instead.
ELs A and B of table 3 report OLS estimates of $\theta_{03}$ and $\theta_{07}$ from regressions with and without boundary fixed effects. Since NCLB scores are measured in logs, their coefficients are elasticities. For example, the results in column 2 indicate that the prices of houses sold in Portland during 2003 were 0.456% higher in attendance zones where NCLB scores were 1% higher. The elasticity is very similar for school districts in Philadelphia (column 3). Notice that Philadelphia is one of four areas to have a significant increase in the price elasticity. It increased from 0.481 in 2003 to 0.710 in 2007 (0.481 + 0.229).

Overall, panel A provides tentative evidence that (i) NCLB scores capture a dimension of school quality that matters for property values and; (ii) the functional relationship between NCLB scores and property values changed over the duration of our study.

The evidence in panel A is tentative because we have not controlled for correlation between school quality and unobserved amenities. Positive correlation seems likely. To see this, first note that household income is a strong predictor of a child’s academic performance.\textsuperscript{33} Now consider a household’s location choice problem. If homebuyers appreciate low crime rates, access to parks, and scenic views, they will bid up prices in neighborhoods with those amenities. Wealthier parents who can afford to live in high-amenity neighborhoods will have children who perform better on standardized tests. Therefore, our inability to control for crime, parks, and views will produce an upward bias on the OLS estimator for the test score coefficient. Boundary fixed effects can mitigate this problem by absorbing the price effect of unobserved amenities in each boundary zone.

\textsuperscript{33} Correlation between household income and academic performance reflects a web of interaction between several underlying factors. Income is correlated with parental education and ability which, in turn, may help to explain the quality of the early parenting environment. Income is also correlated with the education and ability of the parents’ of the child’s peers, and so on. While positive correlation between income and test scores is sufficient to develop intuition for the endogeneity problem in our model, understanding the underlying causal mechanisms is critical to the development of effective education policies. See Heckman (2008) for a summary of the evidence.
Panel B reports regression results after adding boundary fixed effects. Consistent with intuition, the coefficients of variation increase and the test score coefficients decrease. Comparing panels A and B reveals that boundary fixed effects decrease most of the elasticities by more than 50%!

NCLB scores are not directly comparable across states because each state develops its own tests. Nevertheless, since the state-specific scores represent different proxy measures of the same underlying variable—school quality—they can be compared in terms of a common proportionate change. The elasticities in columns 6-10 are remarkably similar across the five metro areas in 2003. They suggest a 1% increase in math and reading proficiency would increase property values by 0.12% to 0.27%. In comparison, Black (1999) reports an increase of 0.42% for Boston in 1993-1995 and the results from Bayer, Ferreira, and McMillan (2007) indicate an increase of 0.12% for San Francisco in 1990.

In 2007 our range of point estimates for the test score elasticity is wider: 0.04 to 0.57. The changes are large and significant for Fairfax, Portland, Detroit, and Los Angeles. Large changes in the test score coefficients signal that the hedonic gradient changed. Moreover, changes in other coefficients are large enough to reject the hypothesis of a time-constant gradient for every metro area (F-tests are reported in panel B). Philadelphia is the only area with a p-value near the 0.05 threshold. These results clearly indicate the pres-

---

34 The impact on the test score coefficients of including the boundary fixed effects is quite similar (in percentage terms) to the results reported by Black (1999) and Bayer, Ferreira, and McMillan (2007). Coefficients on the control variables are generally consistent across metro areas with the usual signs and plausible magnitudes. Results are suppressed for brevity and will be provided upon request. Like Bayer, Ferreira, and McMillan we find that, more often than not, inclusion of the boundary fixed effects decreases the magnitudes of the coefficients on neighborhood demographics.

35 Several factors may be contributing to these changes, including: (i) changes in NCLB scores; (ii) changes in wealth; (iii) the information shock created by the format for tracking performance under the NCLB program; (iv) changes in neighborhood demographics; (v) changes in other housing characteristics that serve as substitutes or complements for school quality; and (vi) changes in the stock of housing. Parsing out the relative importance of these effects would require estimating the demand curve for school quality. See Epple (1987) and Ekeland, Heckman, and Nesheim (2004) for an overview of the challenges with hedonic demand estimation.
ence of conflation bias.

6.2. Capitalization Effects Measured Over 5-Year Intervals

To assess the magnitude of conflation bias, we use capitalization effects to calculate another set of test score elasticities. We regress price changes on changes in test scores, treating the average house in each block group as an observation.\textsuperscript{36} The control variables include changes in tax rates, changes in the characteristics of residents living in each block group, and changes in the physical characteristics of the average house sold within the block group. Differencing the data purges omitted characteristics of block groups that are constant between 2003 and 2007.

Panel C of table 3 reports results based on the full sample of block groups. Los Angeles is the only area where the capitalization effect implies an elasticity (0.17\%) within the range defined by the parameters of the 2003 and 2007 price functions (0.14\% to 0.22\%). In Fairfax, Portland, Philadelphia, and Detroit, the capitalization effects are far below the lower bound of point estimates from single year price functions. The implied elasticity is at least positive and marginally significant in Philadelphia. In Fairfax and Portland, the elasticities are close to zero. In Detroit the estimated capitalization effect is negative and marginally significant. This could reflect specification error in the linear form of the estimating equation, but the hedonic estimates in column 9 seem plausible by contrast.

The results in panel C may simply be confounded by omitted variables. As we noted earlier, schools with lower test scores in 2003 tended to experience larger increases in test

\textsuperscript{36} As noted earlier, there are too few repeated sales of individual houses to support a micro data analysis. Our use of block group averages provides greater resolution than recent studies that defined the unit of observation as a census tract median or a county average (Chay and Greenstone 2005, Greenstone and Gallagher 2008, Baum-Snow and Marion 2009).
scores. These increases may not be exogenous. If changes in unobserved attributes of block groups are negatively correlated with changes in scores, then our estimators for capitalization effects may be biased downward. Recall that we control for changes in property taxes and changes in observable characteristics of block group populations. This means that any confounding must come from changes in unobserved variables that co-vary with changes in NCLB scores across the block groups within a metro area (e.g. crime rates). While this is certainly possible, it seems unlikely that localized amenities (other than school quality) would change sharply as we cross an attendance zone boundary. Based on this logic, adding boundary fixed effects to the regression will mitigate potential confounding by absorbing the capitalization of changes in all unobserved variables that are common to both sides of a 0.2-mile boundary zone.

To implement our panel data version of the boundary discontinuity design we first drop all houses located more than 0.2 miles from a boundary. Then we aggregate the micro data on either side of each boundary. Finally, we add fixed effects for each boundary zone and estimate the resulting first-differenced model,

\[
\Delta \ln P = \Delta \text{test score}\phi + \Delta h\gamma + BFE + \Delta \varepsilon.
\]

This specification uses the same geography and the same fixed effects as the cross-section regressions. If changes in omitted variables are negatively correlated with changes in test scores, then we would expect the capitalization effects to increase. Results are reported in panel D. While standard errors on the capitalization-based elasticities have increased due to the decrease in sample size and the inclusion of fixed effects, the point estimates are very similar to our baseline results. The estimates in columns 16-20 all fall within 95%
confidence intervals of the estimates from columns 11-15. Thus, we do not find strong evidence of confounding from time-varying omitted variables.

We ran additional robustness checks to evaluate the scope for aggregation bias (i.e. using block groups to estimate capitalization effects) and sample selection bias (i.e. using 0.2 mile boundary zones to estimate price functions) but did not find evidence to suggest that either issue can explain the differences between our estimates for capitalization effects and our estimates for price function parameters. Details are provided in the appendix.

6.3. Implications for Welfare Measurement

The results from our single-year regressions suggest that hedonic price functions adjusted to changes in housing market conditions. These changes matter for evaluating the benefits of public education. Table 4 provides a summary comparison between hedonic and capitalization based estimates for the average resident’s willingness to pay for a 1% increase in NCLB scores. Each column reports the MWTP predicted by a specific model, averaged over the samples from all five study regions. In columns 1-3 we do not control for omitted variables. The resulting predictions are fairly robust to how we define a data point (house, block group) and how we define the extent of the market (full metro area, 0.2 mile boundary zone). However, these predictions are twice as large as predictions from models using boundary fixed effects to mitigate confounding (column 4).

The boundary discontinuity design in column 4 is our preferred specification. It mitigates confounding in a way that is widely believed to be credible (e.g. see Greenstone and Gallagher 2008 p.997); it controls for race-based sorting (Bayer, Ferreira, and McMillan 2007); and it predicts MWTP using data for a single school year, which seems consistent
with the static description of equilibrium in Rosen’s model. It implies the average household would have been willing to pay $536 (year 2000 dollars) for a 1% improvement in school quality in 2003. Looking across metro areas, average MWTP ranges from $422 for Detroit to $743 for Philadelphia. This range lies within the range of estimates for San Francisco ($372) and Boston ($917) reported by Black and Bayer, Ferreira, and McMillan.

There were several changes to housing market conditions between 2003 and 2007. Property values increased by 6% on average, test scores increased by 10% on average, and there were smaller changes in the demographic compositions of neighborhoods. There was also steady media coverage of the NCLB program and changes in the broader economy that would have affected expectations about permanent income (e.g. rapid growth in stock market indices and personal income). These changes were accompanied by changes in hedonic gradients which, in turn, increased our prediction for average MWTP to $688 in 2007.

Capitalization effects suggest much smaller measures of MWTP. Column 5 reports the average MWTP predicted by the simple first-differenced model ($134 in 2003, $152 in 2007).37 These figures are about ¼ the size of estimates from single-year boundary discontinuity regressions! The difference only narrows slightly when boundary fixed effects are used to mitigate confounding by time-varying omitted variables (column 6). Overall, these results reinforce the cumulative evidence from our theoretical and econometric models, suggesting that caution is needed in using capitalization effects to calculate welfare measures.

37 These figures were calculated by combining results from columns 11-15 in table 3 with data on average property values and populations in tables 2 and A1-A4.
As in all empirical hedonic studies, our ability to translate price function parameters into measures of MWTP relies on several maintained assumptions. In particular, we have assumed that each metro area represents a fully integrated housing market, and that households are myopic, freely mobile, and fully informed about the spatial distribution of public goods in their home market. These assumptions serve as caveats to our findings and deserve some discussion.\footnote{Modeling forward looking behavior and moving costs has been found to be important for welfare measures estimated from structural models of household sorting (Bayer, Keohane, and Timmins 2009, Bishop and Murphy 2011). The implications for reduced form estimation of hedonic price functions have yet to be investigated.}

What is the spatial extent of an implicit market for a public good? This remains an open question. Some empirical hedonic studies define the market to be national, assuming the price function gradient is stable throughout the country. This assumption would not make sense for our application to school quality because each state develops its own testing system. While we allow the hedonic gradient to vary across metro areas, it would also be interesting to investigate the extent to which hedonic gradients vary within metro areas. The challenge for future research is to develop a research design that would enable one to identify the spatial extent of the market during a single time period, while simultaneously addressing the omitted variable problem.

In order for housing prices to adjust to changes in test scores within a metro area, a significant number of homeowners must move and movers must be informed about test scores. While free mobility is an idealized description of reality, Americans \textit{are} remarkably mobile. Migration statistics from the Census Bureau indicate that, on average, 13.3%
of U.S. residents moved each year between 2003 and 2007. Did perceptions about school quality influence their location choices? Evidence can be found in the 2003, 2005, and 2007 versions of the American Housing Survey. Recent homebuyers were asked to name the main factors that influenced their migration decisions. Their responses do not indicate that changes in test scores triggered them to move out of their old neighborhoods. However, conditional on having decided to move, perceptions about school quality influenced where many movers decided to live. Each year, between 14% and 18% of homebuyers cited “good schools” as one of the reasons they chose to move into their new neighborhoods. A steady 7% cited good schools as the main reason.

Unfortunately, we are unable to document how homebuyers formed their perceptions about school quality. This is a common limitation. Within the hedonic literature, it is standard to assume that buyers make choices based on the same objective measure of a spatially delineated amenity that is collected by the analyst. While this may be a strong assumption for some amenities, it seems relatively plausible in the case of school quality. As we noted earlier, “school quality” was defined in our regressions using the same NCLB proficiency statistics that were distributed to parents and posted on popular websites such as greatschools.org. Thus, our measure of school quality, however imperfect, coincides with information that would have been easily obtained by homebuyers.

Finally, it is important to acknowledge a limitation of our estimated capitalization effects. Recall that we selected school quality for the application in order to track potential changes in single-year price functions using a credible research design for mitigating

See http://www.census.gov/hhes/migration/.

Most migrants moved for personal reasons such as establishing their household, a new job, marriage or divorce.

“Convenient to job” is the most frequently cited reason, followed by “convenient to friends or relatives”.

39
40
41
omitted variable bias. The case for identifying capitalization effects in our first-differenced regressions is admittedly weaker. We did not develop an instrument for the change in test scores. Previous studies that developed instruments for changes in amenities used data throughout the United States over 10 to 20 year periods. Research designs based on national data and 10 to 20 year periods raise serious concerns about time-varying omitted variables. There is less scope for this source of confounding in our application, by definition, since we study much smaller geographic areas over much shorter time periods. Nevertheless, we did not simply dismiss the problem. To mitigate potential confounding, we used boundary fixed effects to absorb the capitalization of unobserved changes common to both sides of each boundary. The resulting loss of precision in our point estimates limits our certainty about the magnitude of conflation bias. However, it does not affect our conclusion that a bias exists.

We observed a large difference between our hedonic and capitalization-based estimates for MWTP, but this large difference is not our main finding. Our main finding is that the evidence from well identified boundary discontinuity regressions strongly suggests that price function gradients changed between 2003 and 2007. These changes violate the time-constant gradient assumption. Thus, the test score capitalization effect does not reveal the public’s marginal willingness to pay for improved test scores. Even the “ideal” instrument for the change in test scores will not be enough to overcome this fundamental economic limitation.

7. Conclusion

Presidential Executive Order 12866 (1993) requires federal agencies to assess “costs
and benefits” of regulations using “the best reasonably obtainable scientific, technical, economic, and other information.” Rosen’s (1974) model provides the starting point for developing revealed preference estimates for the benefits of changes in public goods and externalities. Can we trust these estimates enough to use them in benefit-cost analyses of federal regulations? A recent wave of hedonic modeling has improved the credibility of estimates by refining research designs to mitigate confounding by omitted variables. However, these refinements interact with the revealed preference logic of Rosen’s model to change the interpretation of the identified parameters.

We have extended Rosen’s model to express reduced form estimates for “capitalization effects” for public goods in terms of structural parameters representing household preferences. It is clear that capitalization effects do not generally reveal well-defined measures of consumer welfare. The problem arises from the market-clearing role of the hedonic price function. Following an exogenous shock to the spatial distribution of a public good, the price function must adjust in order to clear the market. We derived the “conflation bias” associated with interpreting the price adjustments as sufficient statistics for benefit measurement, and then we investigated the implications for valuing school quality. There was strong evidence of conflation bias and suggestive evidence that it caused us to understate the willingness to pay for improvements to school quality by as much as 75%.

The lack of a welfare interpretation does not make capitalization effects uninteresting or unimportant. Capitalization effects matter to homeowners, renters, assessors, appraisers, and to the beneficiaries of programs funded by property tax revenue. Credible estimates for capitalization effects may therefore have many practical applications. However, it
would be prudent to avoid interpreting capitalization effects as direct measures for the benefits of public programs, unless that interpretation can be supported by evidence that the gradient of the hedonic price function was stable over the duration of the study period.

REFERENCES


Chetty, Raj. 2009. “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and


TABLE 1
SUMMARY STATISTICS OF SCHOOL TEST SCORE DIFFERENCES

<table>
<thead>
<tr>
<th></th>
<th>FAIRFAX, VA</th>
<th>PORTLAND, OR</th>
<th>PHILADELPHIA, PA</th>
<th>DETROIT, MI</th>
<th>LOS ANGELES, CA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
</tr>
<tr>
<td>2002/2003 math-reading score</td>
<td>81.88 10.44</td>
<td>79.35 11.34</td>
<td>67.43 13.19</td>
<td>67.17 11.39</td>
<td>45.73 17.34</td>
</tr>
</tbody>
</table>

Changes in math-reading score

<table>
<thead>
<tr>
<th></th>
<th>FAIRFAX, VA</th>
<th>PORTLAND, OR</th>
<th>PHILADELPHIA, PA</th>
<th>DETROIT, MI</th>
<th>LOS ANGELES, CA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
<td>mean  sd</td>
</tr>
<tr>
<td>10th decile</td>
<td>11.35 8.03</td>
<td>1.45 9.22</td>
<td>18.78 0.80</td>
<td>15.08 2.43</td>
<td>10.60 3.26</td>
</tr>
<tr>
<td>middle deciles</td>
<td>0.62 5.02</td>
<td>-4.02 6.61</td>
<td>10.45 4.85</td>
<td>11.15 2.78</td>
<td>9.24 2.01</td>
</tr>
<tr>
<td>90th decile</td>
<td>-0.44 2.79</td>
<td>-4.50 4.07</td>
<td>6.28 1.36</td>
<td>5.13 1.72</td>
<td>5.97 1.51</td>
</tr>
</tbody>
</table>

% change in 10th decile | 13.87% | 1.83% | 27.85% | 22.46% | 23.18%

NOTE.—Means and standard deviations for test scores are based on NCLB information aggregated and reported by [www.schooldatadirect.org](http://www.schooldatadirect.org). The math reading score is an overall measure (calculated by Standard & Poor’s) that provides an average of the proficiency rates achieved across all reading and math tests, weighted by the number of tests taken for each elementary school (Fairfax and Portland) or school district (Philly, Detroit and LA). Raw scores are not directly comparable across states because each state develops its own standardized tests.
## TABLE 2
**SUMMARY STATISTICS FOR HOUSING, NEIGHBORHOODS, AND TEST SCORES IN FAIRFAX, VA**

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Full Sample (micro data: N = 10,255)</th>
<th>Sample: 0.20 Mile Boundary Zone (micro data: N = 5,843)</th>
<th>Full Sample (Census block group data: N = 438)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (1)</td>
<td>standard deviation (2)</td>
<td>mean (3)</td>
</tr>
<tr>
<td>Fairfax County, VA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sale price</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 price</td>
<td>567,322</td>
<td>247,727</td>
<td>546,575</td>
</tr>
<tr>
<td>2007 price</td>
<td>562,683</td>
<td>305,748</td>
<td>542,998</td>
</tr>
<tr>
<td>Average math/reading test result</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 score</td>
<td>83.56</td>
<td>9.54</td>
<td>83.01</td>
</tr>
<tr>
<td>2007 score</td>
<td>84.36</td>
<td>8.25</td>
<td>83.90</td>
</tr>
<tr>
<td>Housing characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>square feet (100's)</td>
<td>21.12</td>
<td>9.93</td>
<td>20.66</td>
</tr>
<tr>
<td>bathrooms</td>
<td>3.24</td>
<td>1.08</td>
<td>3.21</td>
</tr>
<tr>
<td>age</td>
<td>34.07</td>
<td>15.82</td>
<td>34.13</td>
</tr>
<tr>
<td>lot acres</td>
<td>0.38</td>
<td>0.43</td>
<td>0.35</td>
</tr>
<tr>
<td>bedrooms</td>
<td>3.94</td>
<td>0.77</td>
<td>3.93</td>
</tr>
<tr>
<td>Neighborhood characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% block group nonwhite</td>
<td>0.23</td>
<td>0.11</td>
<td>0.23</td>
</tr>
<tr>
<td>% block group under 18</td>
<td>0.24</td>
<td>0.04</td>
<td>0.24</td>
</tr>
<tr>
<td>% block group owner occupied</td>
<td>0.85</td>
<td>0.15</td>
<td>0.84</td>
</tr>
<tr>
<td>% block group vacant</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>block group pop density</td>
<td>0.37</td>
<td>0.22</td>
<td>0.40</td>
</tr>
<tr>
<td>tax rate</td>
<td>111.85</td>
<td>49.52</td>
<td>111.45</td>
</tr>
</tbody>
</table>

**NOTE.**—This table reports summary statistics for the key variables included in the analysis for Fairfax, VA. Cols. 1, 2, 3, 6 and 7 are simply the means and standard deviations for the 3 different samples of data. The boundary zone sample includes all houses located within 0.20 miles of the boundary of another school attendance zone. Col. 4 reports the difference in means between houses located on the “high” test score side of a boundary with the corresponding mean for the “low” test score houses on the opposite side of the boundary. Col. 5 provides a T-statistic on the difference in these means. Cols. 8 and 9 report correlations between the change in test scores and levels and changes in all other variables for the full sample of census block group data.
### TABLE 3
TEST SCORE COEFFICIENTS FROM HEDONIC AND CAPITALIZATION REGRESSIONS

<table>
<thead>
<tr>
<th></th>
<th>FAIRFAX, VA</th>
<th>PORTLAND, OR</th>
<th>PHILADELPHIA, PA</th>
<th>DETROIT, MI</th>
<th>LOS ANGELES, CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Test Score Parameters from Hedonic Regressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(micro data from 0.2 mile sample without boundary fixed effects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log (test score), 2003 coefficient</td>
<td>0.122</td>
<td>0.456</td>
<td>0.481</td>
<td>0.524</td>
<td>0.274</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.020)</td>
<td>(0.045)</td>
<td>(0.036)</td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>log (test score), 2007 differential</td>
<td>0.554</td>
<td>0.034</td>
<td>0.229</td>
<td>0.516</td>
<td>0.084</td>
</tr>
<tr>
<td>(0.056)</td>
<td>(0.032)</td>
<td>(0.067)</td>
<td>(0.086)</td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.74</td>
<td>0.70</td>
<td>0.68</td>
<td>0.68</td>
<td>0.75</td>
</tr>
<tr>
<td>Number of observations</td>
<td>6,036</td>
<td>14,443</td>
<td>3,973</td>
<td>6,252</td>
<td>12,287</td>
</tr>
<tr>
<td>B. Test Score Parameters from Hedonic Regressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(micro data from 0.2 mile sample with boundary fixed effects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log (test score), 2003 coefficient</td>
<td>0.116</td>
<td>0.200</td>
<td>0.272</td>
<td>0.208</td>
<td>0.140</td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.028)</td>
<td>(0.071)</td>
<td>(0.047)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>log (test score), 2007 differential</td>
<td>0.293</td>
<td>-0.165</td>
<td>-0.120</td>
<td>0.357</td>
<td>0.075</td>
</tr>
<tr>
<td>(0.081)</td>
<td>(0.048)</td>
<td>(0.101)</td>
<td>(0.126)</td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.85</td>
<td>0.77</td>
<td>0.76</td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td>Number of observations</td>
<td>6,036</td>
<td>14,443</td>
<td>3,973</td>
<td>6,252</td>
<td>12,287</td>
</tr>
<tr>
<td>F-test on H₀: time-constant gradient</td>
<td>4.69</td>
<td>1.98</td>
<td>1.86</td>
<td>4.41</td>
<td>8.22</td>
</tr>
<tr>
<td>p-value on F-test</td>
<td>0.000</td>
<td>0.031</td>
<td>0.047</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>C. Test Score Parameters from Capitalization Regressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(block group data from full sample without boundary fixed effects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>change in log (test score)</td>
<td>-0.037</td>
<td>0.007</td>
<td>0.116</td>
<td>-0.289</td>
<td>0.174</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.096)</td>
<td>(0.068)</td>
<td>(0.134)</td>
<td>(0.033)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.53</td>
<td>0.45</td>
<td>0.29</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Number of observations</td>
<td>438</td>
<td>754</td>
<td>1,199</td>
<td>1,477</td>
<td>6,975</td>
</tr>
<tr>
<td>D. Test Score Parameters from Capitalization Regressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(block group data from 0.2 mile sample with boundary fixed effects)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>change in log (test score)</td>
<td>0.008</td>
<td>-0.025</td>
<td>0.130</td>
<td>-0.445</td>
<td>0.231</td>
</tr>
<tr>
<td>(0.111)</td>
<td>(0.091)</td>
<td>(0.180)</td>
<td>(0.521)</td>
<td>(0.177)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.83</td>
<td>0.82</td>
<td>0.91</td>
<td>0.87</td>
<td>0.83</td>
</tr>
<tr>
<td>Number of observations</td>
<td>422</td>
<td>603</td>
<td>176</td>
<td>213</td>
<td>251</td>
</tr>
</tbody>
</table>

NOTE.—All regressions included controls for property taxes, structural housing characteristics (square feet, number of bathrooms, age, lot size, number of bedrooms) and neighborhood characteristics measured at the block group level (population density, percent nonwhite, percent under 18, percent owner occupied, and percent vacant). In cols. 1 through 10, the dependent variable is the natural log of the sale price of the house. All control variables are interacted with a dummy for sales made during the 2007-2008 school year. In cols. 11 through 20 the dependent variable is the change in the natural log of the average sale price in the census block group. All regressions use Eicker-White standard errors.
<table>
<thead>
<tr>
<th>Identification strategy:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>full</td>
<td>full</td>
<td>0.2 mile</td>
<td>0.2 mile</td>
<td>full</td>
<td>0.2 mile</td>
</tr>
<tr>
<td>Data point</td>
<td>block group</td>
<td>house</td>
<td>house</td>
<td>house</td>
<td>block group</td>
<td>block group</td>
</tr>
<tr>
<td>Sample size</td>
<td>23,149</td>
<td>244,551</td>
<td>42,991</td>
<td>42,991</td>
<td>10,843</td>
<td>1,665</td>
</tr>
<tr>
<td>Controls for omitted variables</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>boundary fixed effects</td>
<td>differencing</td>
<td>boundary fixed effects</td>
</tr>
</tbody>
</table>

NOTE.—All measures of willingness to pay are reported in constant year 2000 dollars. Each measure is averaged over the samples from our five study regions, using the elasticities reported in tables 3 and A1. For example, the estimates in col. 4 are based on the elasticities reported in cols. 6 through 10 of table 3.
SUPPLEMENTAL APPENDIX: FOR ONLINE PUBLICATION ONLY

Section I uses Tinbergen’s “linear-quadratic-normal” model to provide an example of the precise relationship between measures of MWTP derived from preference functions and capitalization effects derived from the adjustments to the reduced form of the equilibrium that clear the market following an exogenous shock. Section II describes robustness checks on the empirical model for sample selection bias and aggregation bias. Section III presents summary statistics for the data describing Portland, OR; Los Angeles, CA; Philadelphia, PA; and Detroit, MI.

I. Consequences of Violating Assumption 1: A Linear-Quadratic-Normal Example

Suppose the housing stock is fixed, utility is quadratic, and heterogeneous preferences and housing characteristics are normally distributed. These assumptions conveniently provide a closed-form linear expression for the equilibrium price function. Specifically, let the utility from a house defined by attributes \( k = \left[ g, X \right] \) be parameterized as

\[
U = -\left( k - \alpha \right)^\prime \Omega \left( k - \alpha \right) + b,
\]

where \( \Omega \) is a positive definite diagonal scaling matrix. When \( k \) and \( \alpha \) are both normally distributed such that \( k \sim N(\mu_k, \Sigma_k) \) and \( \alpha \sim N(\mu_\alpha, \Sigma_\alpha) \), the price function can be expressed as

\[
P(k) = \Psi \left( \alpha(k) \right) + \frac{1}{2} \Gamma \left( \alpha(k) \right), \quad \text{where} \quad \Psi = \Omega \left( \mu_\alpha, -\Sigma_\alpha^{0.5} \Sigma_k^{0.5} \mu_k \right) \quad \text{and} \quad \Gamma = -\Omega \left( I - \Sigma_\alpha^{0.5} \Sigma_k^{0.5} \right).
\]

Notice that the reduced-form parameters \((\Psi, \Gamma)\) are functions of the structural parameters that describe the distributions of household preferences \((\mu_\alpha, \Sigma_\alpha)\) and housing characteristics \((\mu_k, \Sigma_k)\). This simple model violates the last two conditions of our assumption 1.

Now consider a shock to \( g \). Before the shock, \( MWTP_1 = \Psi_1 + \Gamma_1 k \). After the shock, \( MWTP_2 = \Psi_2 + \Gamma_2 k \). It follows from (A2) that, in general, \( \Psi_1 \neq \Psi_2 \) and \( \Gamma_1 \neq \Gamma_2 \). The rate at which the shock is capitalized into property values is

\[
\phi = \frac{P_2 \left( \cdot \right) - P_1 \left( \cdot \right)}{g_2 - g_1} = \frac{\Psi_2 k_2 + \frac{1}{2} \Gamma_2 k_2 - \Psi_1 k_1 - \frac{1}{2} \Gamma_1 k_1}{g_2 - g_1}.
\]

The difference between \( MWTP_1 \), \( MWTP_2 \), and approximations based on \( \phi \) will depend on the correlation between \( g_1 \) and \( \Delta g \). We demonstrate this numerically, using the following values
for the structural parameters,

\[(A4)\]

\[
\begin{align*}
 k &= [g, x_1, x_2], \\
 \mu_x &= \begin{bmatrix} 20 & 50 & 25 \end{bmatrix}, \\
 \mu_k &= [5, 10, 0], \\
 \Sigma_x &= \begin{bmatrix} 2 & 0 & 0 \\
 0 & 1 & 0 \\
 0 & 0 & 3 \end{bmatrix}, \\
 \Omega &= \begin{bmatrix} 1 & 0 & 0 \\
 0 & 2 & 0 \\
 0 & 0 & 3 \end{bmatrix}, \\
 \Sigma_k &= \begin{bmatrix} 2 & 0 & 0 \\
 0 & 1 & 0 \\
 0 & 0 & 1 \end{bmatrix}.
\end{align*}
\]

These parameter values imply that all three characteristics are normal goods and the demand for each is downward sloping. The multivariate normal distributions \((A4)\) are used to take one million random draws. After evaluating the price function in period 1, we introduce a large shock to \(g\) using \(\Delta g \sim N(3,0.25)\) and \(\text{cov}(\Delta g, x_1) = \text{cov}(\Delta g, x_2) = 0\). Then we evaluate the price function in period 2 and determine the capitalization effect.

Figure A1 reports results for two different values of \(\text{cov}(\Delta g, g_1)\). Each panel shows the implicit marginal price functions for \(g\) before and after the shock, as well as demand curves for two households. Because demand is downward sloping, a positive shock increases the price of housing but decreases the MWTP for a further improvement.

![Figure A1](image_url)

**Panel A**: Gentrification

**Panel B**: Preferential attachment

**FIGURE A1.**—Difference between capitalization effects and MWTP.
In panel A, $g_1$ and $\Delta g$ are negatively correlated so that areas with the lowest baseline levels of $g$ receive the largest improvements. This results in a sufficiently large upward bias on the capitalization-based estimate for MWTP ($18.12), that it exceeds the true average MWTP in the pre-shock equilibrium ($15). Intuitively, this is gentrification. The households who value $g$ the most bid up prices in improved areas by more than the average resident is willing to pay.

Panel B demonstrates the opposite case where areas with the highest baseline levels of $g$ receive the largest improvements. In this case, the capitalization-based estimate for MWTP ($6.03) understates average MWTP in the post-shock equilibrium ($11.85). This example of preferential attachment is consistent with Starrett’s (1981) observation that there is little upward pressure on prices when the highest quality neighborhood is improved. Together, panels A and B illustrate how the market sorting process that underlies a hedonic equilibrium can drive a wedge between capitalization effects and MWTP.

![Figure A2](image)

$\text{cov}(\Delta g, g_1) = 0.0$

- mean pre-shock MWTP: $15.00$
- mean post-shock MWTP: $11.92$
- capitalization effect $12.08$

**FIGURE A2.**—Difference between capitalization effects and MWTP.

Finally, figure A2 reports results for a case where $\text{cov}(\Delta g, g_1) = 0$. This setup mimics a randomized experiment. $\Delta g$ is independent of initial levels and changes in all other variables.
While the marginal price function shifts, the capitalization effect ($12.08) provides an excellent approximation to average MWTP in the new equilibrium ($11.92). The econometrics behind this result are explained in section III of the paper.

II. Robustness Checks for Sample Selection Bias and Aggregation Bias

In addition to conflation bias and time-varying omitted variables, two other features of our research design seem capable of explaining the large differences between our baseline estimates for capitalization effects and hedonic price function parameters: sample selection and data aggregation. First consider the scope for sample selection. Houses located outside the 0.2 mile boundary zones are included in the capitalization model but excluded from the hedonic regressions. The excluded houses comprise a large share of total housing sales in each metro area, from 35% in Portland to 92% in Los Angeles. Differences between the capitalization and hedonic results could arise from differences in the distribution of properties located in the excluded and included areas. To evaluate this possibility, we repeat estimation of the basic hedonic model (without boundary fixed effects) using all of the micro data that were used to construct the block group averages for the capitalization model. Results are reported in columns 1-5 of table A1. They essentially mirror the original hedonic estimates from columns 1-5 of table 3. Given the large sample sizes, it is remarkable that only two of the ten coefficients are statistically different (Fairfax and Detroit in 2003). Based on these results, we do not see strong evidence that sample selection is driving the differences between the hedonic and capitalization models.

A second possibility is that the capitalization results are driven by aggregation bias that arises from averaging the micro data over Census block groups. The issue is that the “average” house in a given block group need not correspond to any point on the hedonic price surface. It is difficult to predict the direction and magnitude of the resulting bias. Past studies that have used Census aggregates have assumed the bias is sufficiently small to ignore (Chay and Greenstone 2005, Greenstone and Gallagher 2008, Baum-Snow and Marion 2009). To evaluate this assumption, we aggregate the micro data from panel A into block groups and repeat the estimation. Results are reported in panel B. Comparing the two panels reveals that aggregation does not affect the general pattern of results. The magnitudes of the coefficients do change a bit, but the differences are mostly insignificant.
TABLE A1
ROBUSTNESS CHECKS ON TEST SCORE COEFFICIENTS

NOTE.— All regressions included controls for property taxes, structural housing characteristics (square feet, number of bathrooms, age, lot size, number of bedrooms) and neighborhood characteristics measured at the block group level (population density, percent nonwhite, percent under 18, percent owner occupied, and percent vacant). The dependent variable is the natural log of the sale price of the house. All control variables are interacted with a dummy for sales made during the 2007-2008 school year. All regressions use Eicker-White standard errors.

III. Summary Statistics for Portland, Philadelphia, Detroit and Los Angeles

Summary statistics are reported in table A2 through A5.
### TABLE A2
**SUMMARY STATISTICS FOR HOUSING, NEIGHBORHOODS, AND PUBLIC SCHOOLS IN PORTLAND, OR**

<table>
<thead>
<tr>
<th>Portland Metro Area</th>
<th>Sample: 0.20 Mile Boundary Zone</th>
<th>Full Sample (micro data: N = 25,294)</th>
<th>Sample: 0.20 Mile Boundary Zone (micro data: N = 16,539)</th>
<th>Full Sample (Census block group data: N = 754)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (1)</td>
<td>standard deviation (2)</td>
<td>mean (3)</td>
<td>standard deviation (4)</td>
</tr>
<tr>
<td>Sale price</td>
<td>241,875</td>
<td>142,991</td>
<td>237,021</td>
<td>2,664</td>
</tr>
<tr>
<td>2003 price</td>
<td>324,181</td>
<td>173,171</td>
<td>316,236</td>
<td>-4,500</td>
</tr>
<tr>
<td>Sale price</td>
<td>241,875</td>
<td>142,991</td>
<td>237,021</td>
<td>2,664</td>
</tr>
<tr>
<td>2003 price</td>
<td>324,181</td>
<td>173,171</td>
<td>316,236</td>
<td>-4,500</td>
</tr>
<tr>
<td>Average math/reading test result</td>
<td>79.82</td>
<td>10.93</td>
<td>79.89</td>
<td>7.41</td>
</tr>
<tr>
<td>2003 score</td>
<td>76.00</td>
<td>10.83</td>
<td>75.96</td>
<td>4.79</td>
</tr>
<tr>
<td>Housing characteristics:</td>
<td>17.88</td>
<td>7.76</td>
<td>17.75</td>
<td>-0.06</td>
</tr>
<tr>
<td>square feet (100's)</td>
<td>2.22</td>
<td>0.93</td>
<td>2.22</td>
<td>0.04</td>
</tr>
<tr>
<td>bathrooms</td>
<td>39.67</td>
<td>30.16</td>
<td>38.39</td>
<td>-0.56</td>
</tr>
<tr>
<td>age</td>
<td>0.20</td>
<td>0.29</td>
<td>0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>lot acres</td>
<td>3.07</td>
<td>0.94</td>
<td>3.07</td>
<td>-0.01</td>
</tr>
<tr>
<td>bedrooms</td>
<td>0.17</td>
<td>0.10</td>
<td>0.17</td>
<td>-0.01</td>
</tr>
<tr>
<td>Neighborhood characteristics:</td>
<td>0.23</td>
<td>0.04</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>% block group nonwhite</td>
<td>0.66</td>
<td>0.19</td>
<td>0.66</td>
<td>0.00</td>
</tr>
<tr>
<td>% block group under 18</td>
<td>0.05</td>
<td>0.03</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>% block group owner occupied</td>
<td>0.53</td>
<td>0.29</td>
<td>0.55</td>
<td>-0.01</td>
</tr>
<tr>
<td>% block group vacant</td>
<td>54.62</td>
<td>8.02</td>
<td>54.74</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

**NOTE.**—This table reports summary statistics for the key variables included in the analysis for Portland, OR. Cols. 1, 2, 3, 6 and 7 are simply the means and standard deviations for the 3 different samples of data. The boundary zone sample includes all houses located within 0.20 miles of the boundary of another school attendance zone. Col. 4 reports the difference in means between houses located on the “high” test score side of a boundary with the corresponding mean for the “low” test score houses on the opposite side of the boundary. Col. 5 provides a T-statistic on the difference in these means. Cols. 8 and 9 report correlations between the change in test scores and levels and changes in all other variables for the full sample of census block group data.
**TABLE A3**
SUMMARY STATISTICS FOR HOUSING, NEIGHBORHOODS, AND PUBLIC SCHOOLS IN PHILADELPHIA, PA

<table>
<thead>
<tr>
<th></th>
<th>Full Sample (micro data: N = 29,333)</th>
<th>Sample: 0.20 Mile Boundary Zone (micro data: N = 3,973)</th>
<th>Full Sample (Census block group data: N = 1,199)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Philadelphia Metro Area</strong></td>
<td>mean (1) standard deviation (2)</td>
<td>mean (3) standard deviation (4)</td>
<td>difference in means (5) T-statistic on difference in means (6) standard deviation (7)</td>
</tr>
<tr>
<td><strong>Sale price</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 price</td>
<td>295,845 188,924</td>
<td>285,243 32,498</td>
<td>4.73 5.44 30.63 4.21</td>
</tr>
<tr>
<td>2007 price</td>
<td>334,662 221,967</td>
<td>324,197 49,222</td>
<td>5.44 26.75 25.67 2.15</td>
</tr>
<tr>
<td><strong>Average math/reading test result</strong></td>
<td>67.88 13.93</td>
<td>69.43 11.05</td>
<td>30.63 30.56 30.63 30.56</td>
</tr>
<tr>
<td>2003 score</td>
<td>78.61 10.90</td>
<td>79.51 7.20</td>
<td>25.67 25.67 25.67 25.67</td>
</tr>
<tr>
<td><strong>Housing characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>square feet (100's)</td>
<td>20.87 9.48</td>
<td>20.03 1.21</td>
<td>4.21 3.42 3.42 3.42</td>
</tr>
<tr>
<td>bathrooms</td>
<td>2.37 1.00</td>
<td>2.28 0.99</td>
<td>3.02 2.93 2.93 2.93</td>
</tr>
<tr>
<td>age</td>
<td>42.03 27.85</td>
<td>46.32 3.50</td>
<td>4.23 4.23 4.23 4.23</td>
</tr>
<tr>
<td>lot acres</td>
<td>0.49 0.65</td>
<td>0.44 0.02</td>
<td>1.15 1.15 1.15 1.15</td>
</tr>
<tr>
<td>bedrooms</td>
<td>3.38 0.77</td>
<td>3.33 0.07</td>
<td>2.82 2.82 2.82 2.82</td>
</tr>
<tr>
<td><strong>Neighborhood characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% block group nonwhite</td>
<td>0.12 0.14</td>
<td>0.11 -0.01</td>
<td>-1.30 -1.30 -1.30 -1.30</td>
</tr>
<tr>
<td>% block group under 18</td>
<td>0.23 0.04</td>
<td>0.22 0.00</td>
<td>2.54 2.54 2.54 2.54</td>
</tr>
<tr>
<td>% block group owner occupied</td>
<td>0.78 0.18</td>
<td>0.79 0.02</td>
<td>3.13 3.13 3.13 3.13</td>
</tr>
<tr>
<td>% block group vacant</td>
<td>0.03 0.03</td>
<td>0.03 0.00</td>
<td>-3.77 -3.77 -3.77 -3.77</td>
</tr>
<tr>
<td>block group pop density</td>
<td>0.34 0.39</td>
<td>0.36 -0.04</td>
<td>-3.83 -3.83 -3.83 -3.83</td>
</tr>
</tbody>
</table>

**NOTE.**—This table reports summary statistics for the key variables included in the analysis for Philadelphia, PA. Cols. 1, 2, 3, 6 and 7 are simply the means and standard deviations for the 3 different samples of data. The boundary zone sample includes all houses located within 0.20 miles of the boundary of another school attendance zone. Col. 4 reports the difference in means between houses located on the “high” test score side of a boundary with the corresponding mean for the “low” test score houses on the opposite side of the boundary. Col. 5 provides a T-statistic on the difference in these means. Cols. 8 and 9 report correlations between the change in test scores and levels and changes in all other variables for the full sample of census block group data.
### Table A4
**Summary Statistics for Housing, Neighborhoods, and Public Schools in Detroit, MI**

<table>
<thead>
<tr>
<th></th>
<th>Detroit Metro Area</th>
<th>Sample: 0.20 Mile Boundary Zone</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (1)</td>
<td>mean (2)</td>
<td>mean (3)</td>
</tr>
<tr>
<td></td>
<td>standard deviation</td>
<td>standard deviation</td>
<td>standard deviation</td>
</tr>
<tr>
<td>Sale price</td>
<td>0.14</td>
<td>219,857</td>
<td>214,048</td>
</tr>
<tr>
<td></td>
<td></td>
<td>131,658</td>
<td>14,186</td>
</tr>
<tr>
<td>2003 price</td>
<td>219,857</td>
<td>131,658</td>
<td>214,048</td>
</tr>
<tr>
<td>2007 price</td>
<td>166,801</td>
<td>131,839</td>
<td>157,640</td>
</tr>
<tr>
<td>Average math/reading test result</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 score</td>
<td>68.76</td>
<td>12.39</td>
<td>67.91</td>
</tr>
<tr>
<td>2007 score</td>
<td>79.28</td>
<td>10.52</td>
<td>78.51</td>
</tr>
<tr>
<td>Housing characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>square feet (100's)</td>
<td>16.57</td>
<td>7.79</td>
<td>16.01</td>
</tr>
<tr>
<td>bathrooms</td>
<td>2.06</td>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>age</td>
<td>46.06</td>
<td>23.24</td>
<td>46.78</td>
</tr>
<tr>
<td>lot acres</td>
<td>0.36</td>
<td>0.52</td>
<td>0.30</td>
</tr>
<tr>
<td>bedrooms</td>
<td>3.15</td>
<td>0.73</td>
<td>3.11</td>
</tr>
<tr>
<td>Neighborhood characteristics:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% block group nonwhite</td>
<td>0.13</td>
<td>0.18</td>
<td>0.12</td>
</tr>
<tr>
<td>% block group under 18</td>
<td>0.23</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td>% block group owner occupied</td>
<td>0.80</td>
<td>0.18</td>
<td>0.82</td>
</tr>
<tr>
<td>% block group vacant</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>block group pop density</td>
<td>0.40</td>
<td>0.28</td>
<td>0.46</td>
</tr>
<tr>
<td>tax rate</td>
<td>27.09</td>
<td>11.25</td>
<td>25.90</td>
</tr>
</tbody>
</table>

**NOTE.**—This table reports summary statistics for the key variables included in the analysis for Detroit, MI. Cols. 1, 2, 3, 6 and 7 are simply the means and standard deviations for the 3 different samples of data. The boundary zone sample includes all houses located within 0.20 miles of the boundary of another school attendance zone. Col. 4 reports the difference in means between houses located on the “high” test score side of a boundary with the corresponding mean for the “low” test score houses on the opposite side of the boundary. Col. 5 provides a T-statistic on the difference in these means. Cols. 8 and 9 report correlations between the change in test scores and levels and changes in all other variables for the full sample of census block group data.
### TABLE A5
SUMMARY STATISTICS FOR HOUSING, NEIGHBORHOODS, AND PUBLIC SCHOOLS IN LOS ANGELES, CA

<table>
<thead>
<tr>
<th>Los Angeles Metro Area</th>
<th>Full Sample (micro data: N = 146,788)</th>
<th>Sample: 0.20 Mile Boundary Zone (micro data: N = 12,287)</th>
<th>Full Sample (Census block group data: N = 6,975)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (1)</td>
<td>standard deviation (2)</td>
<td>mean (3)</td>
</tr>
<tr>
<td><strong>Sale price</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 price</td>
<td>460,747</td>
<td>353,758</td>
<td>509,207</td>
</tr>
<tr>
<td>2007 price</td>
<td>486,752</td>
<td>463,063</td>
<td>563,566</td>
</tr>
<tr>
<td><strong>Average math/reading test result</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003 score</td>
<td>39.75</td>
<td>13.75</td>
<td>41.86</td>
</tr>
<tr>
<td><strong>Housing characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>square feet (100's)</td>
<td>17.06</td>
<td>7.67</td>
<td>17.11</td>
</tr>
<tr>
<td>bathrooms</td>
<td>2.13</td>
<td>0.86</td>
<td>2.16</td>
</tr>
<tr>
<td>age</td>
<td>43.06</td>
<td>22.98</td>
<td>44.87</td>
</tr>
<tr>
<td>lot acres</td>
<td>0.25</td>
<td>0.38</td>
<td>0.20</td>
</tr>
<tr>
<td>bedrooms</td>
<td>3.18</td>
<td>0.87</td>
<td>3.21</td>
</tr>
<tr>
<td><strong>Neighborhood characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% block group nonwhite</td>
<td>0.26</td>
<td>0.18</td>
<td>0.31</td>
</tr>
<tr>
<td>% block group under 18</td>
<td>0.25</td>
<td>0.06</td>
<td>0.25</td>
</tr>
<tr>
<td>% block group owner occupied</td>
<td>0.68</td>
<td>0.21</td>
<td>0.70</td>
</tr>
<tr>
<td>% block group vacant</td>
<td>0.06</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>block group pop density</td>
<td>0.72</td>
<td>0.60</td>
<td>0.82</td>
</tr>
<tr>
<td>tax rate</td>
<td>84.33</td>
<td>137.66</td>
<td>82.44</td>
</tr>
</tbody>
</table>

**NOTE.**—This table reports summary statistics for the key variables included in the analysis for Los Angeles, CA. Cols. 1, 2, 3, 6 and 7 are simply the means and standard deviations for the 3 different samples of data. The boundary zone sample includes all houses located within 0.20 miles of the boundary of another school attendance zone. Col. 4 reports the difference in means between houses located on the “high” test score side of a boundary with the corresponding mean for the “low” test score houses on the opposite side of the boundary. Col. 5 provides a T-statistic on the difference in these means. Cols. 8 and 9 report correlations between the change in test scores and levels and changes in all other variables for the full sample of census block group data.