



Estimating the marginal willingness to pay function without instrumental variables[☆]

Kelly C. Bishop^a, Christopher Timmins^{b,c,*}

^a Department of Economics, Arizona State University, Box 879801, Tempe, AZ 85287, United States

^b Department of Economics, Duke University, Box 90097, Durham, NC 20778, United States

^c NBER, United States

ARTICLE INFO

JEL classification:

Q50
Q51
R21
R23

Keywords:

Crime
Hedonic demand
Willingness to pay

ABSTRACT

The hedonic model of Rosen (1974) has become a workhorse for valuing the characteristics of differentiated products despite a number of well-documented econometric problems, including a source of endogeneity that has proven difficult to overcome. Here we outline a simple, likelihood-based estimation approach for recovering the marginal willingness-to-pay function that avoids this endogeneity problem. Using this framework, we find that marginal willingness-to-pay to avoid violent crime increases by sixteen cents with each additional incident per 100,000 residents. Accounting for the slope of the marginal willingness-to-pay function has significant impacts on welfare analyses.

1. Introduction

Dating back to the work of Court (1939); Griliches (1961), and Lancaster (1966), hedonic techniques have been used to estimate the implicit prices associated with the attributes of differentiated products. The seminal work of Rosen (1974) proposed a theoretical structure for the hedonic regression and a two-stage procedure for the recovery of marginal willingness-to-pay (MWTP) functions for the characteristics of differentiated products. Importantly, his two-stage approach allowed households' heterogeneous MWTPs to be functions of the quantities of the product attribute that they consume.¹ This is particularly important when considering non-marginal policy changes (i.e., any change that is large enough to alter the household's willingness to pay at the margin). The two-stage procedure suggested by Rosen (and further developed by subsequent authors) uses variation in implicit prices (obtained either by employing data from multiple markets or by allowing

for non-linearity in the hedonic price function) to identify the MWTP function.

With Rosen (1974) as a backdrop, the property value hedonic model has become the workhorse for valuing local public goods and environmental amenities, despite a number of well-known and well-documented econometric problems.² An important problem arises in the second stage of Rosen's two-step procedure. In separate papers, Bartik (1987) and Epple (1987) describe a source of endogeneity that is difficult to overcome using standard exclusion restriction arguments. Specifically, they note that the unobserved determinants of tastes affect both the quantity of an amenity that a household consumes and the hedonic price of the amenity. In a regression like the one described in the second stage of Rosen's two-step procedure, the quantity of the amenity that a household consumes will therefore be endogenous. Moreover, because of the equilibrium features of the hedonic model, there are very few natural exclusion restrictions that one can use to solve this endo-

[☆] We thank Spencer Banzhaf, John Clapp, Dennis Epple, Matt Kahn, Nick Kuminoff, Alvin Murphy, Holger Sieg, and seminar/conference participants at Arizona State University, Iowa State University, London School of Economics, Montana State University, Syracuse University, University of Aberdeen, University of Arizona, University of Stirling, University of Wisconsin, Washington University in St. Louis, AERE Annual Meetings, Camp Resources, Econometric Society Summer Meetings, Heartland Environmental and Resource Economics Workshop, Krak Foundation, Triangle Resource and Environmental Economics Workshop, and UAE Annual Meetings for their helpful comments and suggestions.

* Corresponding author at: Department of Economics, Duke University, Box 90097, Durham, NC 20778, United States

E-mail addresses: kelly.bishop@asu.edu (K.C. Bishop), christopher.timmins@duke.edu (C. Timmins).

¹ Similarly, the MWTP function may be allowed to vary with the level of *quality* of the attribute that they consume.

² See Taylor (2003) and Palmquist (2005) for a comprehensive discussion. Some of these problems arise in the first stage of Rosen's two-step procedure; for example, omitted variables that may be correlated with the local attribute of interest. There is a large and growing literature that describes both quasi-experimental and structural solutions to this problem (see Parmeter and Pope, 2009 for a discussion).

ogeneity problem.³ With a few exceptions, the hedonics literature has subsequently ignored Rosen's second stage, focusing instead on recovering estimates of the hedonic price function and valuing only marginal changes in amenities.^{4,5}

Various instrumental variables strategies have been proposed to deal with this problem. [Epple \(1987\)](#) provides a set of order conditions that describe a valid set of instruments. [Bartik \(1987\)](#) suggests instrumenting for the quantity of an amenity that a household consumes with market indicator variables and [Kahn and Lang \(1988\)](#) suggest a similar instrument of market indicators interacted with household demographic attributes. The intuition for these strategies is that differences in the distribution of suppliers across markets will provide an exogenous source of variation in the equilibrium quantity of the amenity chosen by each household. The concern with these approaches is that they require strong assumptions about cross-market preference homogeneity and the instrument may not induce sufficient variation in the endogenous variable. [Ekeland et al. \(2004\)](#) takes advantage of the non-linearity of the hedonic model to propose an alternative instrumental-variables strategy which does not require assumptions about cross-market preference homogeneity and may be used in a single-market setting.

In this paper, we describe a simple estimation procedure for the recovery of the marginal willingness-to-pay function. Our parametric approach employs insights from the non-parametric estimators developed in [Ekeland et al. \(2004\)](#) and [Heckman et al. \(2005\)](#) and makes explicit the relationship between the quantity of the amenity being consumed and the attributes of the households doing the consumption. That such a relationship should exist in hedonic equilibrium goes back to the idea of "stratification" found in [Ellickson \(1971\)](#), which became the basis for estimable Tiebout sorting models.⁶ This approach works even in a single-market setting, given a flexible representation of the hedonic price function. Moreover, the approach is computationally simple and does not require any more in terms of data or assumptions than does the standard Rosen-style approach.

We implement this estimation procedure using data on large changes in violent crime rates in the San Francisco Bay Area over the period 1994 to 2000. We find that recovering the full MWTP function is economically important; a household's marginal willingness to pay to avoid an incident of violent crime (measured by cases per 100,000 residents) increases by sixteen cents with each additional incident. Non-marginal reductions in crime of the sort seen in San Francisco and the rest of the nation during the 1990s therefore have the potential to significantly affect MWTP. We find that naive estimation approaches, which ignore this effect, yield estimates of total willingness to pay for crime reductions in San Francisco that are significantly biased. The naive models overestimate the willingness to pay for a crime reduction and underestimate the

³ For example, the exclusion restrictions typically used to estimate a demand system (i.e., using supplier attributes as instruments) will not work because of the sorting process underlying the hedonic equilibrium.

⁴ See, for example, [Black \(1999\)](#); [Gayer et al. \(2000\)](#); [Bui and Mayer \(2003\)](#); [Davis \(2004\)](#); [Figlio and Lucas \(2004\)](#); [Chay and Greenstone \(2005\)](#); [Linden and Rockoff \(2008\)](#); [Pope \(2008\)](#); [Greenstone and Gallagher \(2008\)](#); [Bajari et al. \(2012\)](#), and [Gamber-Rabindran and Timmins \(2013\)](#).

⁵ [Deacon et al. \(1998\)](#) noted that "To date no hedonic model with site specific environmental amenities has successfully estimated the second stage marginal willingness to pay function." Since that time, a number of papers have examined the problem of recovering preferences from hedonic estimates. [Bajari and Benkard \(2005\)](#) avoid the Bartik-Epple endogeneity problem by relying on parametric assumptions on utility that turn Rosen's second-stage from an estimation problem into a preference-inversion procedure. [Yinger \(2014\)](#) uses distributional assumptions about latent demand and unobserved heterogeneity, allowing the preference parameters to be recovered from the price function itself. [Ekeland et al. \(2004\)](#) provide an alternative approach to recovering MWTP that imposes very little in terms of parametric restrictions, but requires an additive separability assumption in the MWTP specification. [Heckman et al. \(2010\)](#) illustrate conditions under which non-separable utility functions may be non-parametrically identified and estimated. Even with these insights, the empirical hedonics literature has largely not moved beyond marginal analyses.

⁶ See, for example, [Epple et al. \(1984\)](#); [Epple and Platt \(1998\)](#); [Epple and Romano \(1998\)](#); [Epple and Sieg \(1999\)](#), and [Kuminoff \(2008\)](#)

willingness to pay to avoid a crime increase. Similar problems are likely to arise in other settings where policy changes are not marginal (e.g., air quality, school reform, and hazardous waste site remediation).

This paper proceeds as follows. [Section 2](#) describes the estimation of the model in detail. [Section 3](#) describes the data used in our application: housing transactions data from the San Francisco Bay Area combined with violent and property crime data from the RAND California Database. [Section 4](#) reports the results of applying this estimation approach to these data. [Section 5](#) calculates the welfare effects from the actual non-marginal changes in crime faced by a subset of homeowners and compares these welfare effects with those calculated with some alternative procedures in the existing literature. Finally, [Section 6](#) concludes.

2. A simple estimation approach for the hedonic model

In this section, we outline an econometric approach for the estimation of the hedonic model which avoids the difficult endogeneity problem altogether. In the following sections, we implement this approach in an application of valuing crime reductions in San Francisco's Bay Area.

Beginning with [Rosen \(1974\)](#), the traditional approach has been to equate the implicit price of an amenity Z (from the estimation of the hedonic price function) to its marginal benefit (which is a function of Z) and use the resulting expression as the estimating equation. The majority of the empirical literature following Rosen has retained this framework while proposing corrective strategies to deal with the endogeneity of the amenity Z . While the first-order conditions for hedonic equilibrium provide a set of equations that will hold in equilibrium, nothing requires us to write the estimating equation in this manner. While this representation does provide an intuitive interpretation of utility maximization, it is the "marginal cost equals marginal benefit" econometric specification itself which has created the endogeneity problem that plagued this literature for decades.

In the basic structure of the hedonic model, there is no fundamental endogeneity problem; when choosing how much of the amenity Z to consume, households take the hedonic price function as given and choose Z to maximize utility based on their individual preferences. These preferences are determined by a vector of observed household characteristics, X , and unobserved taste shifters, ν . As ν and X are typically assumed to be orthogonal in the hedonic model, we are left with a familiar econometric modeling environment: an endogenous outcome variable, Z , which is a function of a vector of exogenous variables, X , and an econometric error, ν .⁷

[Ekeland et al. \(2004\)](#) derive an expression for the distribution of Z conditional on X . This expression is then used as the basis for a non-parametric maximum-likelihood estimator in [Heckman et al. \(2005\)](#). Employing the insights of that framework, we show that the parameters of the parametric model may be easily recovered using simple and straightforward estimation techniques. Intuitively, the approach finds the parameters of the MWTP function that maximize the likelihood of observing each household's chosen Z . We consider the general case in which a closed-form solution for Z may not exist.⁸ We show that by using a simple change-of-variables technique it is straightforward to compute

⁷ In our empirical application, Z is exposure to violent crime and X includes income and a set of race dummies.

⁸ The lack of a closed-form solution will be the case for most non-linear gradient specifications (including the non-parametric specification that we estimate in [Section 5](#)). As a general rule, one should not expect the hedonic price gradient to be linear. Additionally, specifying the MWTP to be linear is not inconsistent with a nonlinear price gradient in equilibrium; as clearly demonstrated by [Ekeland et al. \(2004\)](#), a linear MWTP function alone does not imply a linear price gradient. For the equilibrium price gradient to be linear (with demand- and supply-side heterogeneity), it would not only require that consumers have a perfectly linear MWTP function, but also that suppliers have a perfectly linear marginal willingness-to-accept function and that both preferences and profits are distributed exactly normal. A key insight of [Ekeland et al. \(2004\)](#) is that very minor perturbations of any of these conditions will lead to substantial non-linearity in the resulting equilibrium price gradient.

the likelihood of observing Z . Even in this general case, estimation is easy and is reduced to a single-parameter numerical search.⁹

2.1. Model and estimation

We begin by specifying a price function that relates the price of a house to its amenities Z , H , and ϵ . We specify the amenity of interest Z and, for convenience, a vector of control amenities, H . These are housing and neighborhood amenities which may determine both utility and price but which are not the focus of the research question. Importantly, when Z is allowed to be a vector, the division of observable amenities into Z and H is without loss of generality as all observable amenities may be included in Z (and H empty). ϵ denotes unobservable housing and neighborhood amenities.¹⁰ The indices, $i = 1, \dots, N$ and $j = 1, \dots, J$, index households and markets, respectively. This price function is known up to the parameter vector β , which is allowed to vary by market:

$$P = P(Z_{i,j}, H_{i,j}, \epsilon_{i,j}; \beta_j) \tag{1}$$

We specify utility as being a function of amenities Z , H , and ϵ , as well as numeraire consumption, C . Preferences are shifted by observed household attributes, X , and by unobserved household attributes, ν with $\text{std} = \sigma$. Utility is known up to the parameter vector α , which is allowed to vary by market:

$$U = U(Z_{i,j}, H_{i,j}, \epsilon_{i,j}, C_{i,j}, X_{i,j}, \nu_{i,j}; \alpha_j) \tag{2}$$

Normalizing the price of numeraire consumption to one, the budget constraint simply states that expenditure on housing and on the numeraire must not exceed income, I :

$$I_{i,j} \geq P(Z_{i,j}, H_{i,j}, \epsilon_{i,j}; \beta_j) + C_{i,j} \tag{3}$$

Assuming that the budget constraint binds allows us to rewrite utility as:

$$U = U(Z_{i,j}, H_{i,j}, \epsilon_{i,j}, (I_{i,j} - P(Z_{i,j}, H_{i,j}, \epsilon_{i,j}; \beta_j)), X_{i,j}, \nu_{i,j}; \alpha_j) \tag{4}$$

We assume that the price function and utility function are additively separable in Z , H , and ϵ . As previously highlighted, this is without loss of generality for the observable attributes when Z is defined as a vector.

We now specify a version of this model that is standard in the literature with a quasi-linear form of utility such that the resulting MWTP function is linear in the scalar amenity of interest and additively separable in the preference shock. While these assumptions are not required for identification or estimation, the resulting model is the simplest model that allows for an estimable slope of the MWTP function. It is also the lightest in terms of data requirements, as it does not require data on household income.

We parametrize utility as:

$$U = \alpha_{0,j} + \alpha_{1,j}Z_{i,j} + \frac{1}{2}\alpha_2Z_{i,j}^2 + \alpha_{3,j}X_{i,j}Z_{i,j} + \nu_{i,j}Z_{i,j} + g(H_{i,j}, \epsilon_{i,j}) + (I_{i,j} - P(Z_{i,j}, H_{i,j}, \epsilon_{i,j}; \beta_j)) \tag{5}$$

where $g(\cdot)$ is an unrestricted function of H and ϵ . The first-order condition for a household's optimal choice of Z is then given by:

$$\alpha_{1,j} + \alpha_2Z_{i,j} + \alpha_{3,j}X_{i,j} + \nu_{i,j} - P'(Z_{i,j}; \beta_j) = 0 \tag{6}$$

The traditional estimation strategy associated with Rosen is to first isolate $P'(Z; \beta)$ on the left-hand side of Eq. (6), replace $P'(Z; \beta)$ with an estimate from the first-stage price-function regressions, and estimate the

resulting equation in a separate second stage, treating ν as the regression error. Eq. (7) gives this estimating equation, where the right-hand side of (7) is the MWTP function:

$$P'(Z_{i,j}; \beta_j) = \alpha_{1,j} + \alpha_2Z_{i,j} + \alpha_{3,j}X_{i,j} + \nu_{i,j} \tag{7}$$

An alternative, and arguably more natural approach, would be to rearrange the first-order condition to isolate the choice of amenity Z on the left-hand side. However, in the absence of parametric assumptions on the price function, $P(Z; \beta)$, it would not be possible to solve for a closed-form for Z . Despite this issue, it remains clear that the consumption of Z would vary with observable household characteristics, X , unobservable preference shocks, ν , and the parameters of the hedonic price function, β . Thus, the traditional approach of estimating Eq. (7) directly would suffer from endogeneity concerns and the slope of MWTP, α_2 , would be biased upward.¹¹

Although Z may not be isolated on the left in Eq. (7), the implicit function contains all of the necessary information to recover the parameters describing household preferences, $\{\alpha_{1,j}, \alpha_2, \alpha_{3,j}, \sigma\}$, conditional on recovering β in a separate first stage.¹² Importantly, a standard maximum-likelihood estimation may be implemented by isolating ν on the left-hand side of Eq. (7) and employing a change of variables from Z to ν . Separability in ν , either additive or multiplicative, is sufficient for forming the likelihood of observing Z . In this case, the closed-form solution for ν is given by:

$$\nu_{i,j} = P'(Z_{i,j}; \beta_j) - \alpha_{1,j} - \alpha_2Z_{i,j} - \alpha_{3,j}X_{i,j} \tag{8}$$

We make the distributional assumption that $\nu \sim N(0, \sigma^2)$.¹³ While this assumption is not necessary for identification, it simplifies the problem and allows us to estimate the model via Maximum Likelihood.¹⁴ Using a textbook application of a change of variables, it is straightforward to form the likelihood of the observed vector $\{Z_{i,j}\}_{i=1}^N$ as:

$$\Pi_{i=1}^N \ell(\alpha, \sigma; Z_{i,j}, X_{i,j}) \tag{9}$$

where $\ell(\alpha, \sigma; Z_{i,j}, X_{i,j}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\nu_{i,j}(\alpha))^2\right\} \left| \frac{\partial \nu_{i,j}(\alpha)}{\partial Z_{i,j}} \right|$

To implement this maximum-likelihood procedure, the differential equation for Z does not need to be solved; one only needs to calculate the value of ν consistent with the observed values of Z (given α , $P'(Z; \hat{\beta})$, and X) using Eq. (8) and the following determinant of the change-of-variables Jacobian:

$$\left| \frac{\partial \nu_{i,j}(\alpha)}{\partial Z_{i,j}} \right| = |P''(Z_{i,j}; \hat{\beta}_j) - \alpha_2| \tag{10}$$

In practice, finding the vector of parameters, $\{\alpha_{1,j}, \alpha_2, \alpha_{3,j}, \sigma\}$, that maximizes the likelihood is straightforward and is reduced to a one-dimensional search over α_2 by concentrating the likelihood function. For each iteration of the numerical optimization (i.e., for each guess of α_2), the likelihood-maximizing values of $\alpha_{1,j}$ and $\alpha_{3,j}$ are recovered through a least-squares regression of $(P'(Z; \hat{\beta}) - \hat{\alpha}_2Z)$ on market indicators and X (and interactions of market indicators and X). The likelihood-maximizing value of σ^2 is recovered as $\frac{1}{N} \sum_{i=1}^N \nu_{i,j}^2$. This numerical simplicity follows from the assumption that ν is distributed normally: for each guess of α_2 , recovering the remaining parameters from a linear

⁹ In the Appendix, we (i) illustrate specific examples in which a closed-form solution for Z may be found, including an example where the structural parameters may be recovered from a reduced-form OLS regression of Z on X and (ii) provide details on the performance of the estimation approach with a series of Monte Carlo experiments.

¹⁰ See Bajari and Benkard (2005) for a discussion of how to interpret ϵ as the residual from a price regression, in which case ϵ can be treated as an observable amenity in the utility function. Alternatively, ϵ can be interpreted as measurement error in the price function. In this case, ϵ will not affect utility.

¹¹ It is clear that Z will be correlated with ν even with data coming from multiple markets.

¹² Note that with a flexible representation of the price gradient, cross-market restrictions are not needed for identification of the model. See Section 2.2 for a discussion of identification.

¹³ We specify a homoskedastic error but one could allow for heteroskedasticity here.

¹⁴ If the true distribution of ν is not normal then estimation should be interpreted as Quasi Maximum Likelihood with the standard consistency results applying. For estimation approaches without distributional assumptions on ν , see Ekeland et al. (2004) and Heckman et al. (2005).

model via maximum likelihood is well-known to be equivalent to estimating via OLS. In the Appendix, we describe a GMM-estimation procedure which does not require a distributional assumption on v , but which does have a much higher computational burden.

Consider now the Rosen-style regression equation in Eq. (7) and the (incorrect) likelihood that is implicitly maximized in an OLS estimation of it. This likelihood would differ from the likelihood given in Eq. (9) by only the Jacobian term, $\frac{\partial v(\alpha)}{\partial Z}$, which explicitly controls for the correlation between Z and v . Thus, the omission of this term in the Rosen-style estimation is akin to incorrectly assuming that there is zero correlation between Z and v , even though this assumption is ruled out by the model.¹⁵ Therefore, the Rosen-style estimation minimizes the sum of squared residuals while the likelihood-based estimation minimizes the sum of squared residuals plus an adjustment term that “corrects” for the correlation between Z and v .

The model presented thus far is one where Z is a scalar amenity. However, this estimation approach could be easily extended to estimate models that consider a K -dimensional vector of amenities denoted $\mathbf{Z} = [Z^1, \dots, Z^K]$.¹⁶ Assuming that the elements of the vector v are distributed jointly normal with mean zero and variance-covariance matrix Σ , the likelihood may be formed by employing a change of variables without solving the differential equations for \mathbf{Z} .¹⁷ The likelihood is given by:

$$\prod_{i=1}^N \ell(\alpha, \Sigma; \mathbf{Z}_{i,j}, X_i)$$

$$\text{where } \ell(\alpha, \Sigma; \mathbf{Z}_{i,j}, X_i) = (2\pi)^{-\frac{K}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{v'_{i,j} \Sigma^{-1} v_{i,j}}{2}\right\}$$

$$\left| \frac{\partial(v_{i,j}^1, \dots, v_{i,j}^K)}{\partial(Z_{i,j}^1, \dots, Z_{i,j}^K)} \right|$$

and

$$v_{i,j} = \begin{bmatrix} v_{i,j}^1 \\ \vdots \\ v_{i,j}^K \end{bmatrix} = \begin{bmatrix} \frac{\partial P(\mathbf{Z}_{i,j}; \beta)}{\partial Z_{i,j}^1} \\ \vdots \\ \frac{\partial P(\mathbf{Z}_{i,j}; \beta)}{\partial Z_{i,j}^K} \end{bmatrix} - \begin{bmatrix} \alpha_{1,j}^1 \\ \vdots \\ \alpha_{1,j}^K \end{bmatrix} - \begin{bmatrix} \alpha_2^{1,1} & \dots & \alpha_2^{1,K} \\ \vdots & \ddots & \vdots \\ \alpha_2^{K,1} & \dots & \alpha_2^{K,K} \end{bmatrix} \begin{bmatrix} Z_{i,j}^1 \\ \vdots \\ Z_{i,j}^K \end{bmatrix} - \begin{bmatrix} \alpha_{3,j}^1 \\ \vdots \\ \alpha_{3,j}^K \end{bmatrix} X_{i,j} \quad (11)$$

Finally, the model presented thus far is one where the MWTP function is both linear and additively separable in the idiosyncratic shock, v . This estimation approach is also easily extended to richer specifications of the MWTP function. Conditional on having access to income, I , this approach could be used to estimate preferences when utility is not quasi-linear. Likewise, this approach could be used when v is not additively separable, as this approach simply requires that v is either additively or multiplicatively separable.

2.2. Identification

In this paper, we combine two sources of identifying power; we take advantage of the identifying power provided by the non-linearity of the price gradient as well as the identifying power provided by data observed across multiple markets. We begin this section with a formal

statement of single-market identification for the baseline model presented in Section 2.1, which follows entirely from previously-established results. We then discuss the additional identifying power that multi-market data provides for this model. Throughout, we offer an intuitive discussion of which features of the data identify the parameters of the MWTP function with a particular emphasis on the identification of the slope parameter of the MWTP function, α_2 .

Ekeland et al. (2004) formally establishes that a class of hedonic models is identified in a single market, as long as the hedonic price gradient is non-linear. More specifically, the scalar, additively-separable, quasi-linear utility function shown in Eq. 5 (consistent with the linear willingness-to-pay function) falls within their class of identified utility functions.¹⁸ Therefore, as we specify a price function with a non-linear gradient, the necessary rank condition is satisfied and Theorem 2 in Ekeland, Heckman, and Nesheim holds.¹⁹ This is the key identification result of that paper. As the identification of Ekeland, Heckman, and Nesheim guarantees identification for our case, the likelihood principle then ensures that the population objective function for our model has a unique maximizer, i.e., for any parameter vector,

$$(\alpha, \sigma) \neq (\alpha^*, \sigma^*) \Rightarrow \ell(\alpha, \sigma | Z, X) \neq \ell(\alpha^*, \sigma^* | Z, X).$$

The identification of the non-slope parameters is straightforward; the MWTP intercept (α_1), the coefficients on individual characteristics (α_3), and the MWTP variance (σ^2) are identified by the average level of Z , the covariance between Z and X , and the variance of Z , respectively. In a single market, the slope parameter (α_2) will be identified by the covariance between Z and nonlinear functions of X . Intuitively, a non-linear price gradient will cause the distribution of Z to be asymmetric, even if the distribution of MWTP intercepts is symmetric. For example, if the price gradient were convex in Z , low-preference households would consume considerably less Z than average-preference households, while high-preference households would consume only slightly more Z than average-preference households. The extent of this asymmetry will be driven by the underlying slope of the MWTP function (a flatter slope is associated with more asymmetry). Therefore, as X (the observable demographic characteristics) shifts the MWTP intercept in a linear (i.e., symmetric) fashion, but affects consumption of Z in a non-linear (i.e., asymmetric) fashion, the covariance between Z and nonlinear functions of X will identify the slope parameter.

It is worth noting that the identification results of Ekeland et al. (2004) do not require a distributional assumption regarding the taste shock, v . In our case, we make the assumption that v is distributed normally to facilitate simple estimation by maximum likelihood. While not required for identification, a distributional assumption may contribute to the identification of the slope parameter. For example, if the gradient is non-linear then the distribution of Z will not be normal even when v is distributed normally. As the extent of this deviation from normality will be driven by the underlying slope of the MWTP function (a flatter slope is associated with a larger deviation), the deviation aids in the identification of the slope parameter.

In addition to the single-market identification from the price-gradient nonlinearity, the use of multi-market data (with cross-market restrictions) to identify the hedonic model has been well-documented since the 1980s (see, for example, Brown and Rosen (1982) and Mendelsohn (1985)). It is important to note that in the absence of any cross-market restrictions, i.e., when all of the parameters of the MWTP function vary by market, this is equivalent to multiple single-market cases and the multi-market data does not aid in identification. However, in

¹⁵ Note that the correctly-specified likelihood of observing $P(Z, \beta)$ differs from Eq. (9) only by the multiplicative constant $\frac{1}{p'(Z, \beta)}$ and, as such, would have the same argmax.

¹⁶ Multiple amenities may be combined into a single index and used to reduce the dimensionality of \mathbf{Z} . See Sieg et al. (2002) for a theoretical foundation for such an index of amenities.

¹⁷ By assuming that the idiosyncratic shocks are independent and that the price gradient is not a function of other housing amenity choices one could estimate each amenity decision separately. If the price gradient were a function of other amenities one could still estimate a given amenity decision in isolation by assuming that the MWTP functions for other amenities were flat.

¹⁸ As noted in Ekeland et al. (2004), if the MWTP function were linear in the log of the amenity, the model would be identified as long as the hedonic price gradient were nonlinear in the log of the amenity. In Heckman et al. (2010), single-market identification is proven for a more general class of utility functions with non-additive heterogeneity. That paper also discusses the identification of non-parametric models of utility with multi-market data.

¹⁹ When Z is a vector of amenities, the necessary conditions for identification are discussed in Nesheim (2015) and Chernozhukov et al. (2017).

the presence of cross-market restrictions, the use of multi-market data provides a number of additional sources of identifying variation.

The first source of identifying variation is the sensitivity of market-specific averages of Z to differences in price gradients across markets. Under the assumption of a common-across-markets MWTP function, this relationship will identify the slope of the MWTP function. For example, if average consumption of Z is considerably smaller in markets with higher implicit prices of Z , a relatively steep MWTP slope will be identified. Under a more flexible specification with market-specific intercepts in the MWTP function, (e.g., the specification in our empirical application where the intercepts of the MWTP function are allowed to vary across geographic markets), the market-specific, average consumption of Z will identify $\{\alpha_{1,j}\}_{j=1}^J$ and will no longer contribute to the identification of the slope of MWTP.

The second source of identifying variation is the sensitivity of market-specific covariances of Z and X to differences in price gradients across markets. For example, when α_2 is large, the cross-market variation in the covariance of Z and X will be relatively small, all else equal. Under a more flexible specification where the coefficients on X are allowed to vary across markets, cross-market differences in the covariance between Z and X will identify $\{\alpha_{3,j}\}_{j=1}^J$ and will no longer contribute to the identification of the slope of MWTP.

The third source of identifying variation is the sensitivity of market-specific variances of Z to differences in price gradients across markets. Intuitively, the market-specific variances of Z (conditional on X) are determined by the market-specific price gradients. Importantly, as the cross-market differences in these variances will be intensified by the underlying slope of the MWTP function (a flatter slope is associated with a higher degree of cross-market differences in the variance of Z), cross-market differences in the variance of Z will aid in the identification of the slope parameter. Under a more flexible specification where σ is allowed to vary across markets, cross-market differences in the variance of Z will identify $\{\sigma_j\}_{j=1}^J$ and will no longer contribute to the identification of the slope of MWTP.

The final sources of identifying variation are the sensitivity of the market-specific covariances of Z with nonlinear functions of X and the sensitivity of market-specific higher-order moments of Z to differences in the price gradients across markets. Following the intuition laid out for the single-market case, non-linear, market-specific price gradients will result in a nonlinear relationship between Z and X with the market-specific level of this non-linearity decreasing in the underlying slope of the MWTP function. Thus, in the multi-market case, the asymmetry differences across markets will aid in the identification of the slope parameter. A similar logic applies to how the slope of the MWTP function affects variation in market-specific deviations of Z from normality.

3. Data

To demonstrate the easy applicability of this estimation approach, we apply it to valuing the willingness to pay to avoid violent crime in the San Francisco Metropolitan Area over the period 1994 to 2000. Further details and results of this application are discussed in Sections 4 and 5.

In the first stage of estimation, we nonparametrically recover the hedonic price function for each year of our sample. In the second stage of estimation, we recover the structural parameters of the linear MWTP function, allowing MWTP to vary with demographic characteristics. Finally, to demonstrate the policy implications of various hedonic estimators, we consider the non-marginal policy analysis of observed changes in crime for the one year period of 1999 to 2000.

To estimate these specifications, we employ a varied set of data from multiple sources. These data and our sample cuts are discussed below.

3.1. Property transactions data

The real-estate transactions data that we employ cover six counties of the San Francisco Bay Area (Alameda, Contra Costa, Marin, San

Table 1
Summary statistics - property transactions.

| Variable | $n = 351,938$ | |
|---------------------|---------------|------------|
| | Mean | Std. Dev. |
| Price | 353,745.04 | 192,444.00 |
| Age | 30.30 | 22.45 |
| Lot size (sq. ft) | 7,172.98 | 8,119.14 |
| Square footage | 1,716.27 | 672.44 |
| Number of bathrooms | 2.12 | 0.72 |
| Number of bedrooms | 3.21 | 0.90 |
| Property crime rate | 1,744.63 | 666.17 |
| Violent crime rate | 422.53 | 186.03 |

Price is expressed in constant 2000 dollars. The crime; rates are per 100,000 residents.

Francisco, San Mateo, and Santa Clara) over the period 1994 to 2000. This dataset (used under a licensing agreement with DataQuick, Inc.) includes dates, prices, loan amounts, and buyers', sellers', and lenders' names for all transactions. In addition, for the final observed transaction of each single-family house, the dataset includes housing characteristics such as exact street address, square footage, year built, lot size, number of bedrooms, and number of bathrooms.

Additional data cuts are made in order to deal with the fact that Data-Quick only reports housing characteristics at the time of the most recent assessment, but we need to use housing characteristics from all transactions as controls in our hedonic price regressions.²⁰ First, to control for land sales or total re-builds, we drop all transactions where "year built" is missing or later than the observed transaction date. Second, to control for major improvements or degradations, we drop any property with an observed average appreciation or depreciation rate exceeding the county- and year- specific mean price change by more than fifty percentage points (in either direction). Additionally, we drop any property that moves more than forty percentage points (in either direction) between transactions in the overall county- and year- specific distributions of price. We drop transactions where the price is missing or zero and, after using the consumer price index to convert all transaction prices into 2000 dollars, we drop one percent of observations from each tail of the price distribution to minimize the effect of outliers.²¹ As we merge the crime data using the property's geographic coordinates, we drop properties where latitude and longitude are missing. Finally, we drop houses with more than three observed transactions over the seven-year sample or more than one sale within a given year.²² Table 1 reports the summary statistics for the property transactions data.

3.2. Household demographic data

For our demographic characteristics, we use information on the race and income of buyers recorded on mortgage applications and published in accordance with the Home Mortgage Disclosure Act (HMDA) of 1975. The HMDA data also describe the mortgage lender's name, the loan amount, the year, and the property's census tract. As the variables of lender name, loan amount, year, and census tract are also available in the DataQuick data, we are able to merge the buyer characteristics data with the property transactions data using the algorithm described in

²⁰ Approximately fifty-five percent of sales are the only observed sale for the property.

²¹ Results are not sensitive to whether we drop or Winsorize observations.

²² To maintain a fairly homogenous sample (i.e., properties that would be competing in the same housing market), we make the additional sample cuts. We drop houses where "year built" is less than 1850, "lot size" is zero or greater than three acres, "square footage" is zero or is greater than 10,000 square feet, total "number of bathrooms" is greater than ten, total "number of bedrooms" is greater than ten, total "number of rooms" is greater than fifteen, and "number of stories" is greater than three.

Table 2
Summary statistics - buyers (Full- and 1999- Samples).

| Variable | Full sample | | 1999 sample | |
|--------------------|-------------|------------|-------------|------------|
| | n = 204,953 | | n = 37,691 | |
| | Mean | Std. dev. | Mean | Std. dev. |
| Price | 362,383.67 | 192,748.52 | 378,956.60 | 197,668.52 |
| Violent crime rate | 416.78 | 171.10 | 359.53 | 138.95 |
| Income | 112,452.41 | 108,823.27 | 112,274.03 | 90,638.58 |
| White | 0.63 | 0.48 | 0.61 | 0.49 |
| Asian | 0.24 | 0.42 | 0.25 | 0.43 |
| Black | 0.03 | 0.18 | 0.03 | 0.18 |
| Hispanic | 0.10 | 0.30 | 0.11 | 0.31 |

Prices and incomes are expressed in constant 2000 dollars. The violent crime rate is per 100,000 residents.



Fig. 1. Locations of crime-reporting cities within the metro area.

Bayer et al. (2016).²³ We drop all households where either race or income is missing.

Table 2 reports the summary statistics for the sample of buyers in both the full sample and in the restricted 1999-only sample.²⁴ This restricted sample will be used in the policy analysis that demonstrates the implications of valuing non-marginal changes in violent crime rates.

3.3. Violent crime data

The violent crime rate that we employ comes from the RAND California database and is defined as the number of incidents per 100,000 residents.²⁵ We consider violent crime as it is a relatively homogeneous measure of crime. Violent crime is reported for each of the 80 cities in the San Francisco Metropolitan Area for each year of our data. Fig. 1 illustrates the locations of these cities. For our analysis, we impute a violent crime rate for each individual house using an inverse distance-squared weighted average of the crime rate in each city.²⁶ As a control in our hedonic regressions, we also create an analogous measure of property crime rates from the RAND California database.²⁷ To mitigate the

²³ Using this algorithm, we are able to uniquely match approximately sixty-six percent of all housing transactions to buyers in the uncleaned HMDA dataset (prior to cuts on race and income). The characteristics of the final sample of buyers and houses is remarkably close to those found in IPUMS samples. See Bayer et al. (2016) for further discussion.

²⁴ Note that these second-stage summary statistics additionally reflect the fact that we trim one percent from the tails of the nonparametrically-estimated gradients within our estimation routine.

²⁵ In the data, violent crime is defined as “crimes against people, including homicide, forcible rape, robbery, and aggravated assault.”

²⁶ Distance is computed using the Great Circle estimator, geographic coordinates of city centroids, and geographic coordinates of each house.

²⁷ Property crime is defined as “crimes against property, including burglary and motor vehicle theft.” We use the property crime rate as a control in our hedonic estimation and focus attention on violent crimes in our valuation exercise, as violent crimes are less likely to be subject to systematic under-reporting (Gibbons, 2004).

effect of outliers, we drop one percent from each tail of the distribution of violent crime.²⁸ Table 1 provides summary statistics for both violent crime and property crime at the level of the house. Table 2 provides summary statistics for violent crime at the level of the buyer.

We see a significant amount of variation (both across houses and through time) in our key variable of interest, violent crime. Figs. 2 and 3 illustrate the distribution of house-specific violent crime rates and the time-trend of mean house-specific violent crime rates, respectively. The rates are given by incidents per 100,000 local residents. The declining trend observed in the San Francisco Metropolitan Area is consistent with the decreases in violent crime observed in most of the US over the same period.

4. Results

4.1. The hedonic price function

In this subsection, we discuss the results from the estimation of the following hedonic price function:

$$P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t) = H'_{i,t} \beta_t^H + f_t(Z_{i,t}; \beta_t) + \epsilon_{i,t} \quad (12)$$

To identify the causal impact of Z on price, we control for observed housing attributes as well as unobserved neighborhood amenities. In Eq. (12), the vector $H_{i,t}$ includes quadratic functions in each of the housing attributes (number of bedrooms, number of bathrooms, square footage, lot size, and age) as well as a set of neighborhood fixed effects at the level of the Census tract.²⁹ These fixed effects control for fixed amenities at the neighborhood level, which may include such things as characteristics of neighbors or of their houses, local school quality, and/or proximity to shopping, restaurants, highways, or other local amenities. Kuminoff et al. (2010) shows that including spatial fixed-effects is the preferred way to deal with (potentially correlated) unobservables at the neighborhood level. In this application, we are able to control for these neighborhood unobservables at the relatively fine level of the Census tract, while our amenity of interest (the violent crime rate) is specified at the level of the house.³⁰ We note, however, that in the presence of within-tract variation in unobserved amenities (not captured by observable house-level attributes) that is correlated with crime, omitted variable bias is a potential concern.

To mitigate against bias coming from functional-form misspecification, we model $f_t(Z; \beta)$ as a flexible function of Z which we recover through local-polynomial estimation. As markets are defined by years in this application, we now employ the subscript t and estimate Eq. (12) separately for each year.

In order to estimate the function $f_t(Z; \beta)$ using local polynomial methods, we must first control for the variation in price due to other housing characteristics, $H' \beta^H$. Following Robinson (1988), we obtain an estimate of β^H , allowing us to move $H' \beta^H$ to the left side of Eq. (12).³¹ We then specify the function $f_t(Z; \beta)$ as being locally quadratic (around each observed value of Z):^{32,33}

$$P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t) - H'_{i,t} \hat{\beta}_t^H = \beta_{0,i^*,t} + \beta_{1,i^*,t} Z_{i,t} + \beta_{2,i^*,t} Z_{i,t}^2 + \epsilon_{i,t} \quad (13)$$

where i^* highlights the fact that the β coefficients hold locally for each observed value of Z in the data. In practice, estimates of the hedonic gradient, i.e., $\hat{\beta}_{1,i^*,t} + 2\hat{\beta}_{2,i^*,t} Z_{i,t}$, are recovered for all 204,953 observed

²⁸ Results are not sensitive to whether we drop or Winsorize observations.

²⁹ According to the U.S. Census Bureau, Census tracts are “designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions” and contain approximately 4000 residents.

³⁰ There are 789 Census tracts in our data.

³¹ See Clapp (2003) for another example of using the Robinson two-step method to estimate hedonic price functions.

³² A benefit of using a locally quadratic approach here, as opposed to a locally linear one, is that we use the second derivative in our second-stage estimation.

³³ There is a slight abuse of notation in Eq. (13), as $H' \hat{\beta}^H$ will not equal $H' \beta^H$ in finite samples. The error is more accurately described as $\epsilon_{i,t} + (H'_{i,t} \hat{\beta}_t^H - H'_{i,t} \beta_t^H)$.

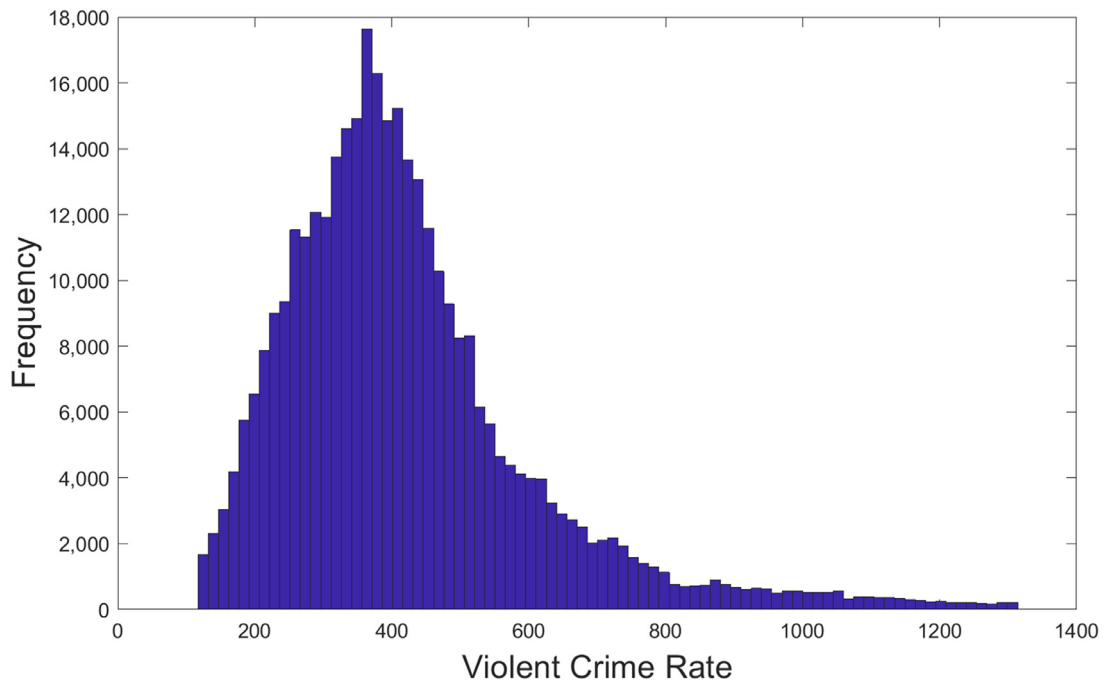


Fig. 2. Overall distribution of violent crime rates.

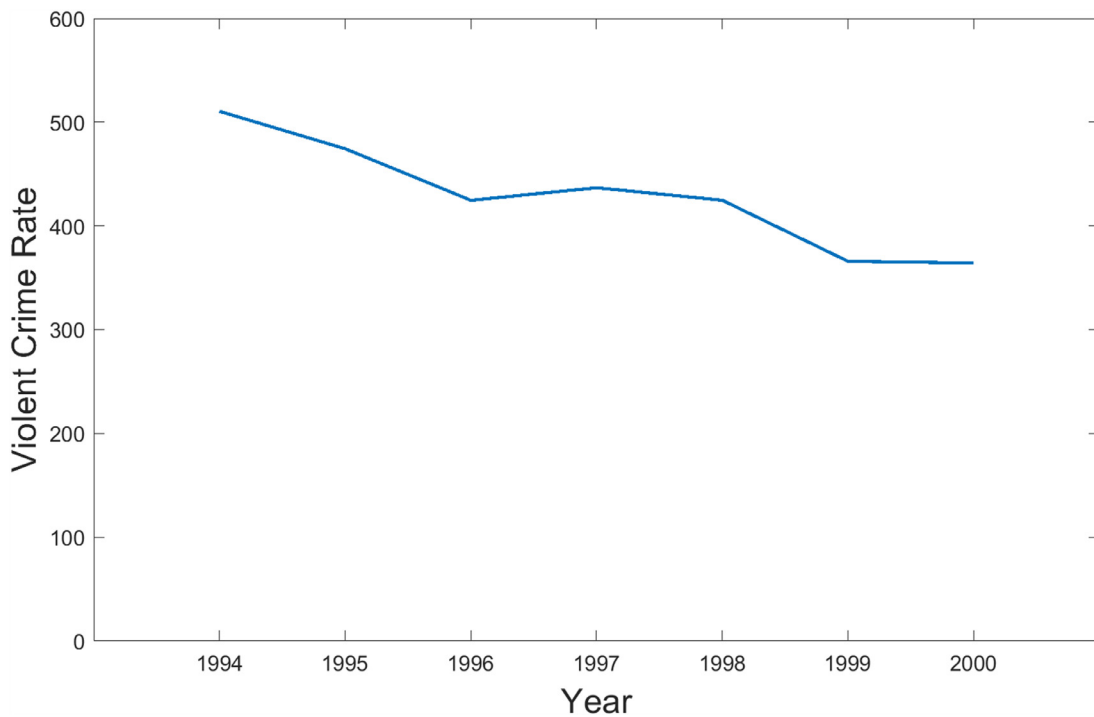


Fig. 3. Time variation in mean violent crime rates.

values of Z in the buyer data. This is in contrast to a fully parametric estimation procedure where a single β would be recovered for a given year.

We estimate Eq. (13) using a Weighted Least Squares routine with weights (as one moves further from the point of evaluation, χ) given by the diagonal of the Gaussian kernel:

$$K_h(Z_{i,t} - \chi) = \frac{1}{h\hat{\sigma}_Z\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{Z_{i,t} - \chi}{h\hat{\sigma}_Z}\right)^2\right\} \quad (14)$$

where h represents the chosen kernel bandwidth and $\hat{\sigma}_Z$ is the standard deviation of Z .

Our selection of bandwidth seeks to allow for maximum flexibility in the function $f_t(\cdot)$, while ensuring that the functions are consistent with utility maximization of homebuyers in the marketplace. In other words, we smooth the function up until the point that all observations exhibit a negative value for the hedonic gradient (the first derivative of price)

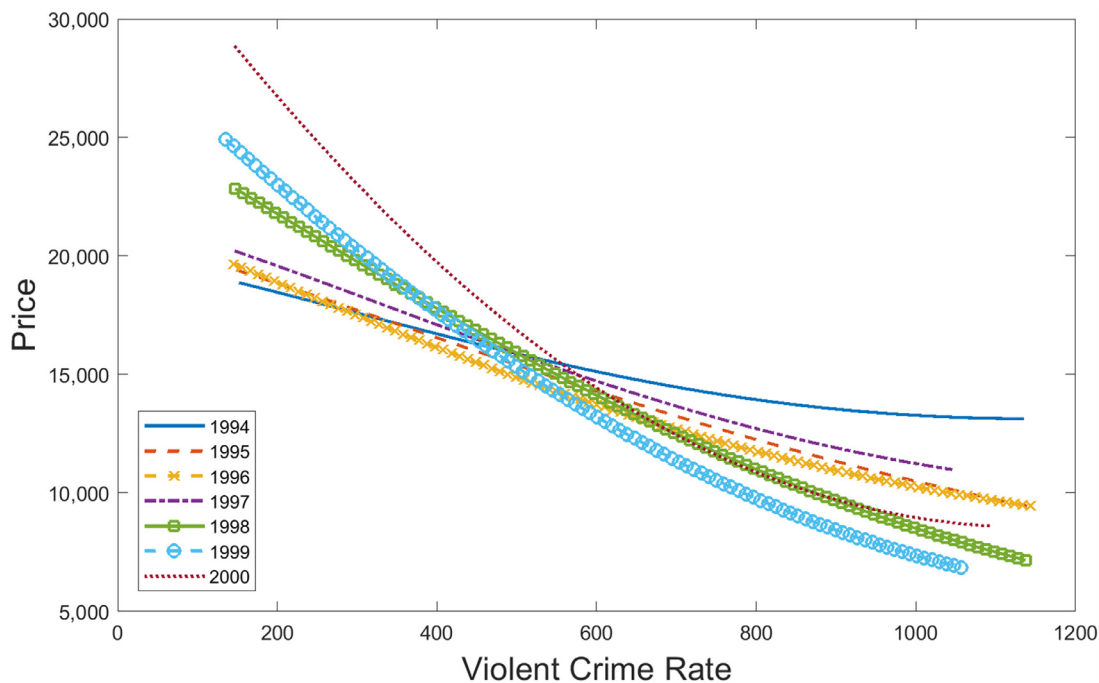


Fig. 4. Results - housing price functions by year, $P_t(Z_{i,t})$.

and a positive value for the second derivatives.³⁴ In practice, we set the bandwidth equal to 2.15 times the standard deviation of violent crime. At this level of smoothing, none of the sample exhibits a positive gradient or a negative second derivative.³⁵ Additional details of Robinson’s estimator may be found in the Appendix.

Results are reported in Figs. 4, 5, and 6, which show the price function, the hedonic gradient (the derivative of price), and the derivative of the gradient (the second derivative of price), respectively.³⁶ Corresponding figures with 99% confidence intervals are shown in Figs. 9, 10, and 11 in the Appendix. The year-specific price functions shown in Fig. 4 plot the rental equivalent of housing price as a function of the violent crime rate.³⁷ Consistent with theory, these functions lie in the first quadrant (positive housing prices) and slope downward (a higher violent crime rate reduces price). The year-specific gradients shown in Fig. 5 plot the derivative of housing price with respect to violent crime as a function of the violent crime rate. The derivative of price can be interpreted as the implicit price of violent crime and is the simple measure of willingness to pay which is often found in the literature. These gradients lie in the fourth quadrant (the implicit price of violent crime is negative or safety is a “good”) and slope upward. Finally, Fig. 6 plots the second derivative of price *w.r.t.* violent crime as a function of the violent crime rate. This figure is useful in seeing both how much heterogeneity exists across years and how much a simple quadratic specification would miss (by forcing these curves to be perfectly horizontal).

³⁴ Note that this second-order condition is stricter than necessary. To be consistent with utility maximization, the second derivative need only be greater than the (negative) slope of the MWTP function.

³⁵ At a tighter bandwidth of 2.1 times the standard deviation of violent crime, there begin to be violations of the second-derivative condition (0.0002 of the sample).

³⁶ For sake of brevity, the β^{H} coefficients (including 788 tract fixed effects), which are not used in estimation, are not reported. They are available upon request.

³⁷ The rental equivalent is taken to be 0.05 of the housing sales price, following the related literature that builds off of Poterba (1984) in expressing housing prices as annualized imputed rents.

4.2. The marginal willingness to pay function

In this subsection, we report the results from the likelihood-based approach for recovering estimates of the MWTP function using the first-stage hedonic price function estimates and the available data on individual home buyers. With the non-parametric specification for the hedonic price function, a closed-form solution for Z cannot be recovered so we employ a change of variables from Z to v . We specify a linear MWTP function and allow the intercept of the MWTP function to vary by market (i.e., by year) by including a set of year dummies in our estimation.³⁸ This implies:

$$P'(Z_{i,t}; \beta_t) = \alpha_{1,t} + \alpha_2 Z_{i,t} + X_i' \alpha_3 + v_{i,t} \tag{15}$$

where:

$$v_{i,t} \sim N(0, \sigma^2) \tag{16}$$

and X_i is a vector comprised of income (in thousands of 2000 dollars) and a vector of race dummies (Asian-Pacific Islander, black, Hispanic, and white).³⁹ As the second stage relies on first-stage estimates, standard errors are bootstrapped by taking 250 replications of the property transactions dataset and estimating both the first and second stages for each replication. An alternative would be to jointly estimate the two stages. While this would increase statistical efficiency, it would greatly increase the computational burden.⁴⁰

³⁸ We can also allow the coefficients on demographic characteristics to vary by year without additional variation or assumptions; employing this specification has almost no impact on results and returns an estimated slope coefficient of -0.1571 . If one were to rely on variation in the price function alone, one could additionally allow the slope coefficient and the error variance to vary by year, returning to the single-market setting.

³⁹ White is the excluded race in our estimation routine.

⁴⁰ To jointly estimate using Maximum Likelihood would also require making a distributional assumption for ϵ .

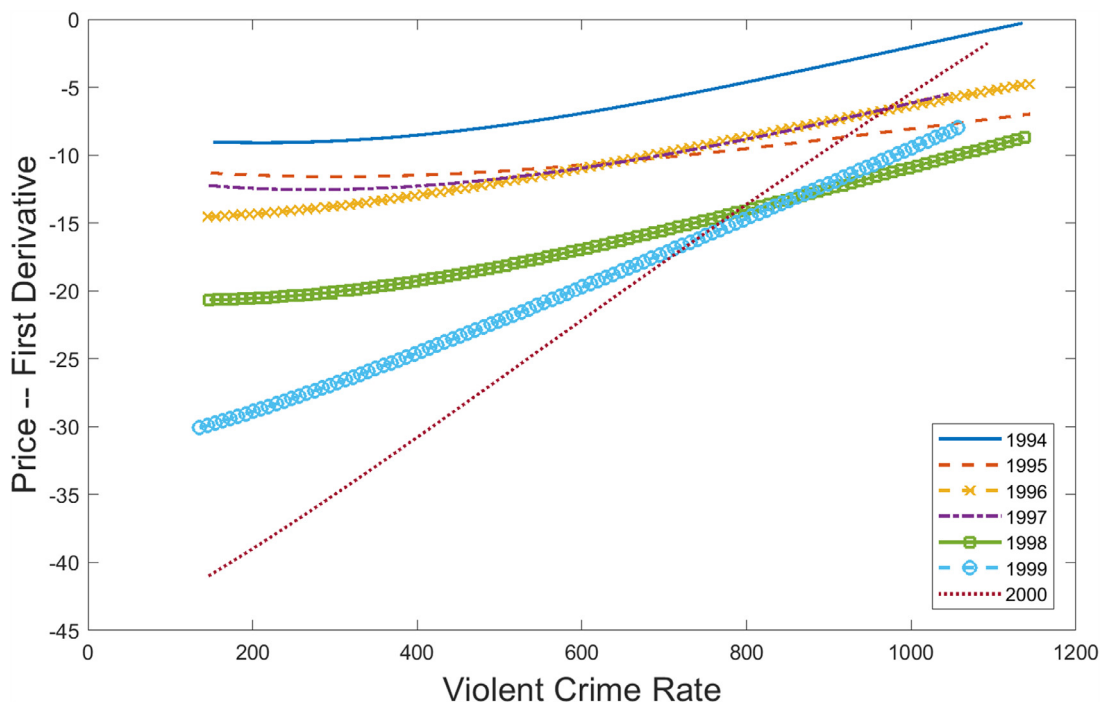


Fig. 5. Results - hedonic gradients by year, $P'_t(Z_{i,t})$.

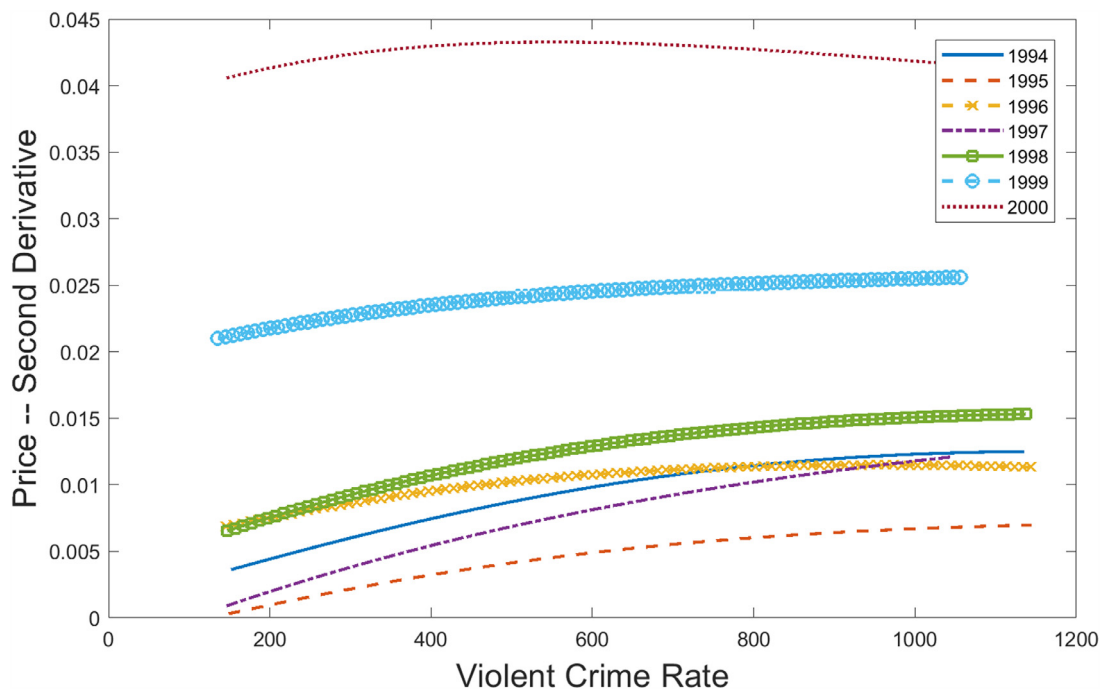


Fig. 6. Results - second derivatives of price by year, $P''_t(Z_{i,t})$.

The first column of Table 3 reports these results.⁴¹ First, consider the coefficient on violent crime, α_2 , which reveals the amount by which the

household’s MWTP to avoid violent crime changes with an increase in this disamenity. Intuitively, this coefficient should be negative, indicating that the MWTP to avoid violent crime increases as the rate of violent

⁴¹ We exclude buyers whose chosen level of crime falls within the first or within the ninety-ninth percentile of violent crime from the first-stage data, as the gradients are estimated non-parametrically and these observations lie close to the boundaries. Importantly,

however, the main parameter of interest (the slope of the MWTP function) only changes from -0.1604 to -0.1793 without these omissions.

Table 3
MWTP function estimates.

| | <i>n</i> = 204,953 | | |
|----------------|--------------------|-------------------|--------------------|
| | Likelihood-based | Rosen | Horizontal |
| Violent Crime | -0.16 (0.02) | 0.02 (0.0008) | – – |
| Income (/1000) | -0.03 (0.005) | -3.37 (0.0001) | -0.003 (0.0003) |
| Asian | 4.09 (0.49) | -0.11 (0.01) | 0.24 (0.03) |
| Black | 24.69 (2.56) | 0.05 (0.04) | 2.10 (0.13) |
| Hispanic | 12.54 (1.35) | 0.06 (0.02) | 1.10 (0.07) |
| Constant | 73.64 (8.93) | -14.79 (0.85) | -7.44 (0.73) |
| 1995 dummy | -9.38 (1.20) | -3.07 (1.05) | -3.59 (1.04) |
| 1996 dummy | -18.69 (1.87) | -3.90 (1.14) | -5.13 (1.14) |
| 1997 dummy | -15.29 (1.45) | -3.27 (0.99) | -4.27 (0.98) |
| 1998 dummy | -24.86 (1.94) | -10.08 (1.16) | -11.31 (1.16) |
| 1999 dummy | -41.25 (2.72) | -15.79 (1.11) | -17.91 (1.11) |
| 2000 dummy | -48.40 (3.36) | -22.85 (1.45) | -24.98 (1.47) |
| σ_v | 27.87 (2.82) | – – | – – |

These are the estimated coefficients from Eq. 7; All are significant at the 1% level of significance.

crime increases (consistent with a demand curve for public safety that is downward sloping). We find this to be the case; each additional incident per 100,000 residents raises MWTP to avoid violent crime by 16.04 cents. As we show in Section 5, this has important implications for the value ascribed to large reductions in violent crime rates like those witnessed over the period of our sample.

Looking at the remaining coefficient estimates, an increase in income of \$1,000 per year increases MWTP to avoid violent crime by 2.95 cents (consistent with public safety being a normal good). Considering differences in MWTP by race, the excluded group (whites) has the highest mean MWTP to avoid violent crime. Results suggest that Asian-Pacific Islanders have a slightly lower mean MWTP (as indicated by their positive intercept shift of \$4.10). Blacks have the lowest mean MWTP to avoid violent crime, followed by Hispanics.

For the sake of comparison, the second column of Table 3 reports the results from the traditional Rosen-estimation approach. These results suggest that increases in violent crime reduce the MWTP to avoid violent crime (indicating that the demand curve for public safety is upward sloping). This is exactly the direction of bias suggested in both Bartik (1987) and Epple (1987) and leads to upwardly-biased estimates of the welfare associated with non-marginal reductions in violent crime (which we show in Section 5).

As the endogeneity bias associated with traditional Rosen approach has been well-documented in the literature, a more natural comparison of our results may be with the Horizontal MWTP. This approach assigns a constant willingness to pay to each household equal to the slope of the hedonic price function at their observed housing choice. As the utility function is assumed to be linear in Z (i.e., the MWTP function is constant in Z), this approach trivially solves the endogeneity problem by assuming that the MWTP function doesn't depend on Z . To do this, we simply run the Rosen second-stage regression restricting $\alpha_2 = 0$. These results are shown in the third column of Table 3. These results are strikingly different from those presented in the first column. While the coefficients on income and race are of the same sign, their magnitudes are much smaller. This indicates that income and race do not appear to play an

economically significant role in these estimates. The economic implications of assuming a horizontal willingness-to-pay will become clear in the analyses in Section 5.⁴²

5. Measuring the welfare implications of a non-marginal change in violent crime rates

As is clear in our data description, the San Francisco Metropolitan Area experienced large and persistent reductions in average violent crime rates over the course of our sample period. Similar reductions have been observed in numerous other cities across the US. Out of the 25 cities that he considers, Levitt (2004) ranks San Francisco 12th in terms of the size of the reduction in homicides experienced between 1991 and 2001. This change in the violent crime rate represents a significant improvement on average and is, importantly, non-marginal.⁴³

There is a large and growing literature aimed at valuing the benefits of crime reductions (with, for example, the goal of conducting cost-benefit analyses of police-force expansions). This literature was recently surveyed by Heaton (2010). He notes that the property value hedonic technique is valuable for recovering the intangible costs of crime (e.g., lost quality of life for fear of victimization or effective loss of public space). Such intangibles are likely to be particularly important for measuring the costs of violent crime (a point emphasized by Linden and Rockoff (2008) with respect to sexual offenses).

For the welfare analyses that follow, we consider the set of all households who purchased a house in 1999 (summary statistics describing these households are given in Table 2). We then measure the value of the crime changes that these households actually experienced between 1999 and 2000. It is highly likely that these households still occupy the same residence in 2000 and, as we show in this section, the changes that occurred over this year were substantial enough that proper identification of the MWTP function becomes important in measuring their change in welfare. In particular, Fig. 7 illustrates the distribution of changes in violent crime rates experienced by this set of households. The 1999 data allow us to consider welfare changes associated with both reductions and increases in crime; approximately two thirds of these households experienced reductions in violent crime during this period while the remaining one third experienced increases in violent crime.⁴⁴

5.1. Welfare analysis using our main specification

In Table 4, we report valuations based on the model presented in Section 2 and the three different estimation strategies discussed in Section 4: (i) the likelihood-based approach presented in this paper, (ii) Rosen, and (iii) Horizontal MWTP. The value associated with a change in the violent crime rate is calculated as the area of under the MWTP function over the width of the experienced change in violent crime. For the Likelihood-based and Rosen strategies, this area will approximate a right trapezoid. For the Horizontal MWTP approach, this area will be a rectangle. See Fig. 8 for illustrations. We report results separately for

⁴² Note that it is difficult to compare goodness of fit across the three approaches, as the Rosen and Horizontal approaches seek to fit the price gradient, $P(Z)$, while our approach seeks to fit the choice of amenity, Z . Furthermore, all the models fail to yield closed-form solutions for Z , which makes constructing a standard R^2 infeasible. However, one can calculate the likelihood of observing the data given the parameters for all three models. As such, we can calculate a McFadden Pseudo R^2 equal to $1 - \frac{\log \ell(\alpha, \sigma)}{\log \ell(\text{intercept})}$. The resulting measures for the Likelihood-based, Rosen, and Horizontal are 0.1983, 0.1478, and 0.1375, respectively.

⁴³ Levitt (2004) discusses six factors that he argues were not responsible for these declines, including economic growth and reduced unemployment, shifting age and racial demographics, changes in policing strategies, changes in gun control laws and laws controlling concealed weapons, and changes in capital punishment. He argues instead that there is a strong case to be made for the role of increasing size of the police force, increased incarceration rates, declines in the crack epidemic, and the legalization of abortion twenty years prior. The relative importance of each of these factors is still a contentious topic. See, for example, Blumstein and Wallman (2006).

⁴⁴ The range of one-year changes in crime is -302.4770 to +616.5721. Only 0.0097 of households experience a change in crime greater than 250 in absolute value.

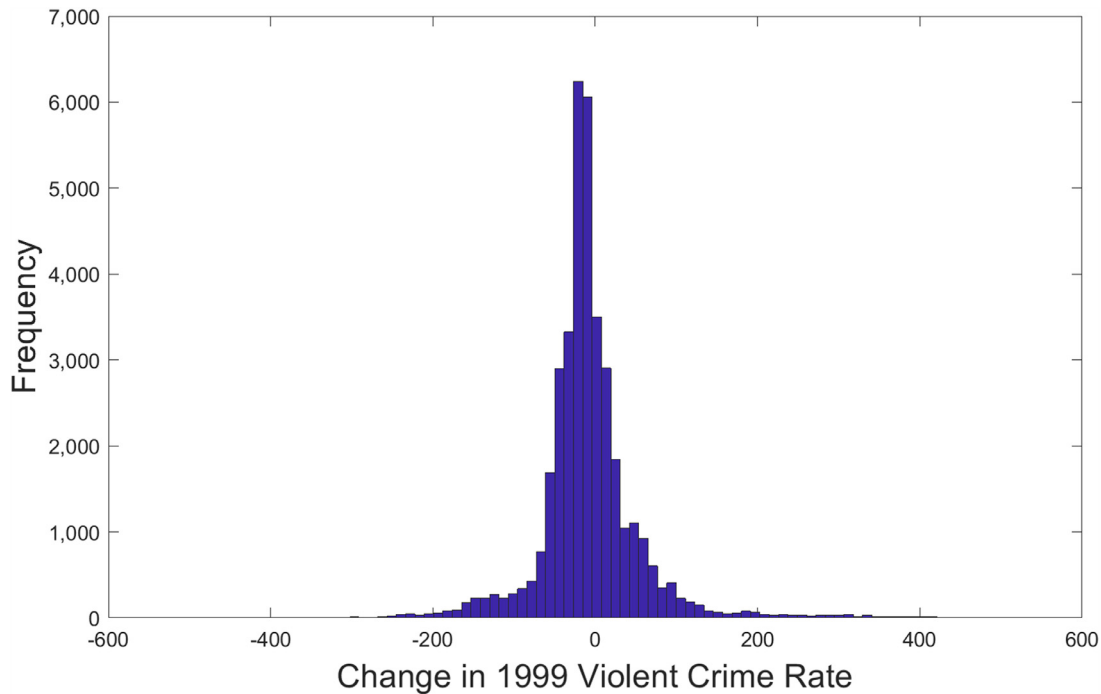


Fig. 7. Distribution of one-year crime rate changes for 1999 buyers.

Table 4
WTP for non-marginal changes in violent crime.

| | Buyers with reductions | | | Buyers with increases | | |
|------------------|------------------------|------------|------------|-----------------------|------------|------------|
| | (n = 24, 791) | | | (n = 12, 900) | | |
| | Average WTP | 25th % WTP | 75th % WTP | Average WTP | 25th % WTP | 75th % WTP |
| Likelihood based | 652 | 294 | 917 | -1710 | -1828 | -290 |
| Rosen | 878 | 309 | 1113 | -1147 | -1501 | -281 |
| Horizontal | 858 | 308 | 1099 | -1194 | -1537 | -282 |

These are estimates of willingness to pay (in year-2000 dollars) for the observed 1999; changes in crime. Welfare estimates are constructed using estimates of the model presented; in Eq. 5.

the two-thirds of buyers who experienced a reduction in their violent crime rate and for the remaining one-third of buyers who experienced an increase in Table 11.⁴⁵

The bias from improperly accounting for the effect of a non-marginal change in violent crime on MWTP is evident. Consider first the case of crime reductions. The Rosen estimation yields estimates of the average WTP for observed reductions that are 1.35 times greater than our model, implying a thirty-five percent upward bias. Results of the Horizontal specification are similar; this specification yields estimates that are 1.32 times greater, implying a thirty-two percent upward bias. The direction of the bias is reversed when we consider increases in the rates of violent crime. Here, the alternative estimation approaches yield estimates of average WTP for observed crime reductions that are only 0.67 and 0.70, respectively, of our Likelihood-Based estimate. These differences are far from trivial and would have important impacts on any cost-benefit analysis.

5.2. Welfare analyses under alternative models

5.2.1. Allowing for non-quasilinear utility

In Section 2.1, we present a quasi-linear model where numeraire consumption, C, enters the utility function in a linear and additively-

separable manner. A benefit of this specification is that data on income are not required for estimation.⁴⁶ Here we present an alternative model where numeraire consumption enters utility non-linearly. Specifically, it enters with a quadratic and with an interaction with Z, allowing both the intercept and the slope of the MWTP function to vary with household budget constraints. We specify household i's utility in year t as:

$$U = \alpha_{0,t} + \alpha_{1,t}Z_{i,t} + \frac{1}{2}\alpha_2Z_{i,t}^2 + \alpha_3X_{i,t}Z_{i,t} + v_{i,t}Z_{i,t} + g(H_{i,t}, \epsilon_{i,t}) + (I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t)) + \alpha_4Z_{i,t}(I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t)) + \frac{1}{2}\alpha_5(I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t))^2 \tag{17}$$

which yields the following first-order condition for Z:

$$\alpha_{1,t} + \alpha_2Z_{i,t} + \alpha_3X_{i,t} + v_{i,t} + \alpha_4(I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t)) - P'(Z_{i,t}; \beta_t)(1 + \alpha_4Z_{i,t} + \alpha_5(I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t))) = 0 \tag{18}$$

Note that this specification yields a closed-form solution for v and the likelihood may be formed using a change of variables from Z to v as described in Section 2.⁴⁷ Likelihood-based estimation results of this

⁴⁵ In the calculations for buyers experiencing a decrease in violent crime, we restrict the marginal willingness to pay curve to be non-positive. This has little effect on the welfare implications; the average willingness to pay without this restriction is 638.02.

⁴⁶ We do allow income to be one of the variables in X. However, this method of incorporating income in the MWTP function is not derived from first principles.

⁴⁷ In this case, finding the vector of parameters that maximizes the likelihood is still straightforward, but is now reduced to a three-dimensional numerical optimization problem (versus the single parameter search described in Section 2.1). Specifically, concentrat-

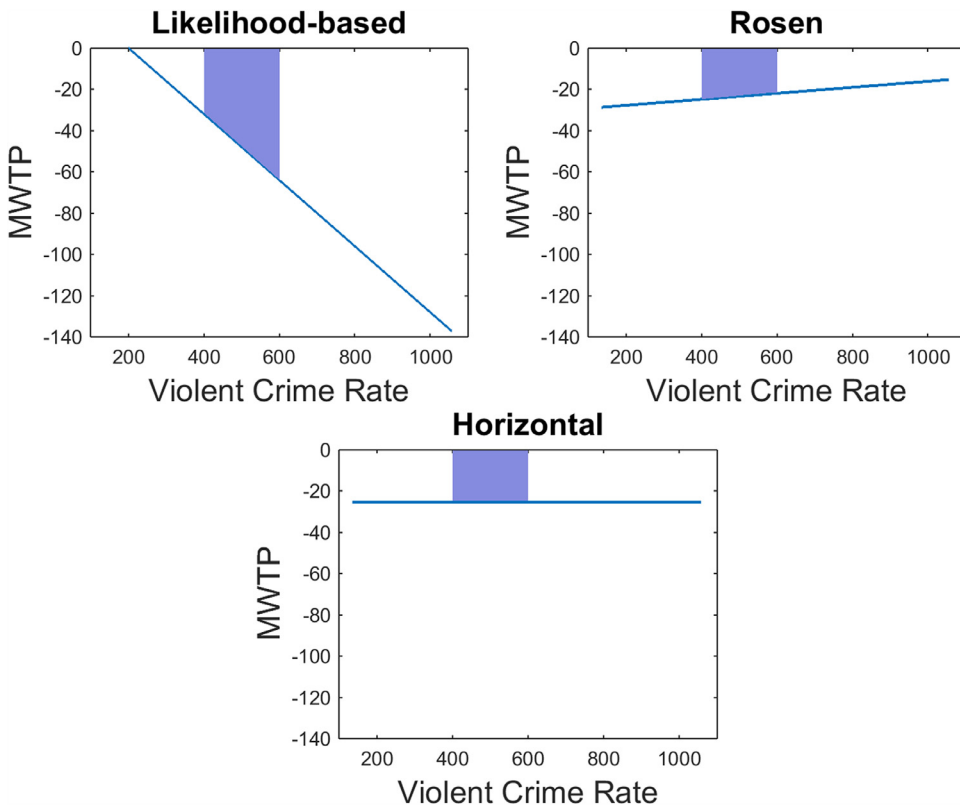


Fig. 8. Illustrations of WTP for non-marginal changes in violent crime.

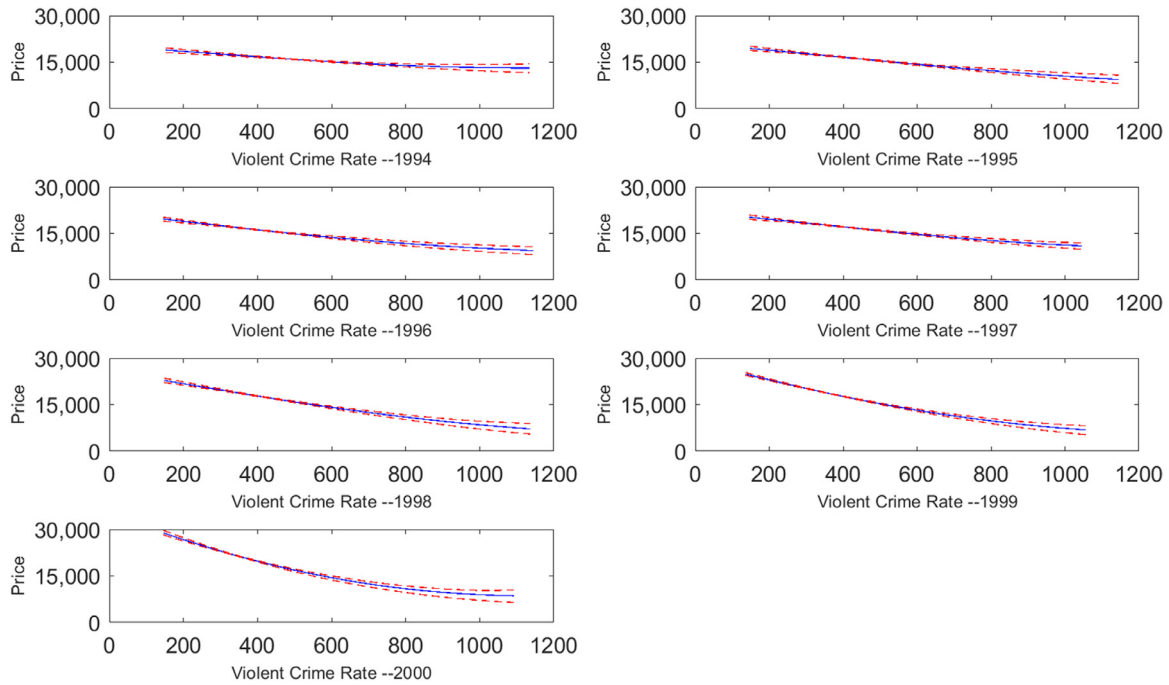


Fig. 9. Results - price functions by year, $P_t(Z)$.

Non-Quasilinear Utility model may be found in Table 5 alongside the corresponding estimates from a Rosen estimation and a Horizontal estimation of this same model.

ing the likelihood function allows for a search over α_2 , α_4 and α_5 where, for each guess of α_2 , α_4 and α_5 , the likelihood-maximizing values of the remaining α parameters are recovered through a least-squares regression and the likelihood-maximizing value of σ^2 is recovered as $\frac{1}{N} \sum_{i=1}^N v_{i,t}^2$.

As before, we report results separately for buyers with reductions and for buyers with increases. By definition, the Horizontal estimates are identical to those in Table 4, as the slope is set to zero in both cases. The results from the Rosen-based estimation approach (which simply minimizes the sum of squared residuals, $\sum_{i=1}^N v_{i,t}^2$) are almost identical to those in Table 4 and still imply an upward-sloping MWTP function.

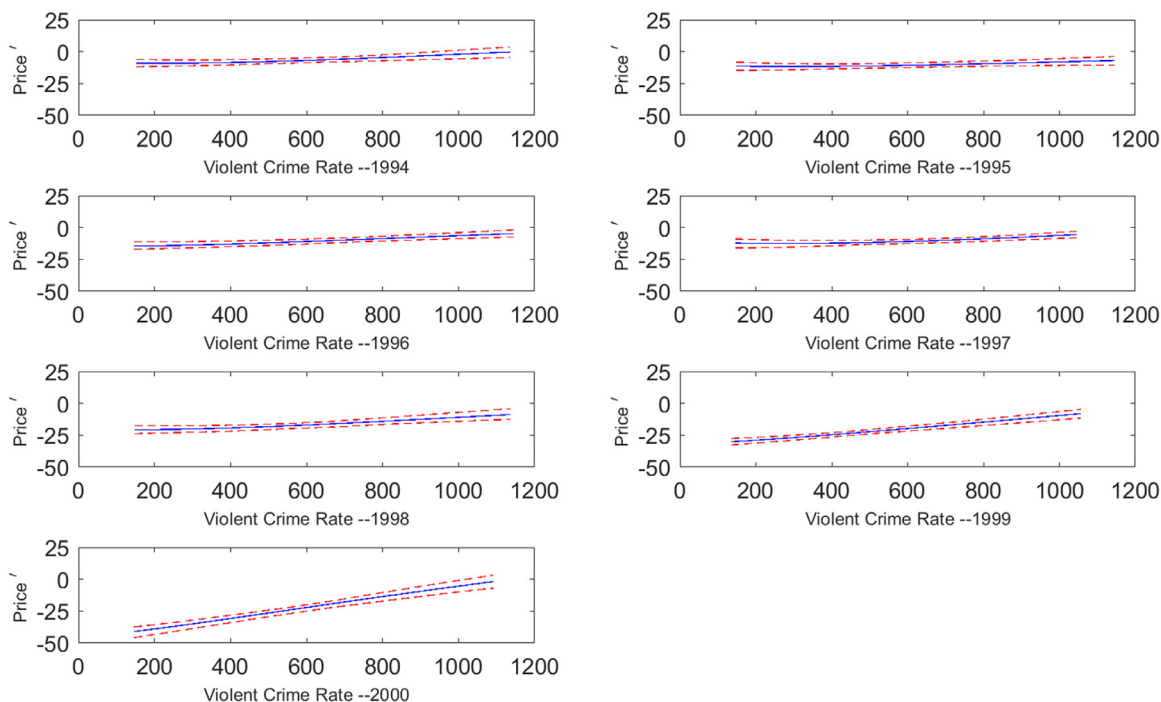


Fig. 10. Results - hedonic gradients by year, $P'_t(Z)$.

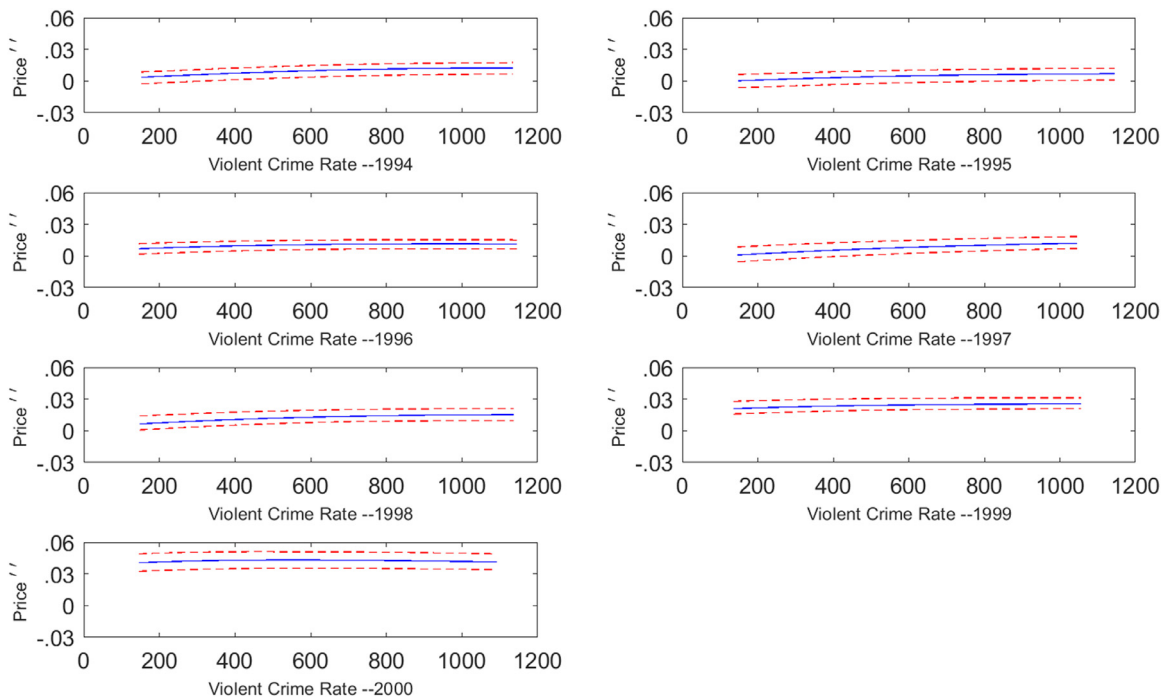


Fig. 11. Results - second derivatives of price by year, $P''_t(Z)$.

The Likelihood-Based results imply a downward-sloping MWTP function with similar welfare effects to those reported in Table 4.⁴⁸

⁴⁸ Not surprisingly, the estimated income elasticities are different. The mean and median income elasticities in the baseline model, which simply allows income to shift the MWTP intercept, are 0.23 and 0.17, respectively. The corresponding elasticities from the Non-Quasilinear Utility model, which allows income to affect behavior through the budget constraint, are 0.51 and 0.38, respectively.

5.2.2. Allowing for non-separable preference heterogeneity in MWTP

In our main empirical specification, the unobservable preference shock, v , enters the MWTP function in an additively-separable manner. Here we present an alternative model in which MWTP is multiplicatively separable in v by specifying household i 's utility in year t as:⁴⁹

⁴⁹ This utility function is analyzed in Heckman et al. (2010).

Table 5
WTP for non-marginal changes in violent crime with non-quasilinear utility.

| | Buyers with reductions | | | Buyers with increases | | |
|------------------|------------------------|------------|------------|-----------------------|------------|------------|
| | (n = 24, 791) | | | (n = 12, 900) | | |
| | Average WTP | 25th % WTP | 75th % WTP | Average WTP | 25th % WTP | 75th % WTP |
| Likelihood based | 634 | 291 | 900 | -1780 | -1853 | -291 |
| Rosen | 878 | 309 | 1113 | -1148 | -1502 | -281 |
| Horizontal | 858 | 308 | 1099 | -1194 | -1537 | -282 |

These are estimates of willingness to pay (in year-2000 dollars) for the observed 1999; changes in crime. Welfare estimates are constructed using estimates of the model presented in Eq. 17.

Table 6
WTP for non-marginal changes in violent crime from the model with non-separable preference heterogeneity.

| | Buyers with reductions | | | Buyers with increases | | |
|------------------|------------------------|------------|------------|-----------------------|------------|------------|
| | (n = 24, 791) | | | (n = 12, 900) | | |
| | Average WTP | 25th % WTP | 75th % WTP | Average WTP | 25th % WTP | 75th % WTP |
| Likelihood based | 648 | 290 | 877 | -2276 | -2250 | -303 |
| Rosen | 893 | 311 | 1128 | -1125 | -1472 | -279 |
| Horizontal | 858 | 308 | 1099 | -1194 | -1537 | -282 |

These are estimates of willingness to pay (in year-2000 dollars) for the observed 1999; changes in crime. Welfare estimates are constructed using estimates of the model presented in Eq. 19.

$$U = \frac{-1}{\alpha_2 + 1} Z_{i,t}^{\alpha_2 + 1} e^{(\alpha_{1,t} + \alpha_3 X_{i,t})} v_{i,t} + g(H_{i,t}, \epsilon_{i,t}) + (I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t)) \tag{19}$$

which yields the following first-order condition for Z:

$$-Z_{i,t}^{\alpha_2} e^{(\alpha_{1,t} + \alpha_3 X_{i,t})} v_{i,t} - P'(Z_{i,t}; \beta_t) = 0 \tag{20}$$

where v is no longer additively separable. This specification yields a constant elasticity of α_2 in the MWTP function, as opposed to our main empirical specification which yields a constant slope.

Note that our likelihood-based approach requires only separability in v ; the preference shock may still be isolated in Eq. 20 and the likelihood may be formed using a change of variables from Z to v .⁵⁰ Likelihood-based estimation results of this Non-Separable Preference Heterogeneity model may be found in Table 6 alongside the corresponding estimates from a Rosen estimation and a Horizontal estimation of this model.⁵¹

As before, the Horizontal estimates are identical to those in Table 4, as the slope is set to zero in both cases. The results from the Rosen-based estimation approach (which simply minimizes the sum of squared residuals, $\sum_{i=1}^N \log(v_{i,t})^2$) are similar to those in Table 4 and still imply an upward-sloping MWTP function. The Likelihood-Based results imply a downward-sloping MWTP function with similar welfare effects to those reported in Table 4 for buyers with reductions. The absolute value of WTP is larger for buyers with increases under Non-Separable Preference Heterogeneity, due to the fact that the constant elasticity leads to a

⁵⁰ Given this form of utility, an obvious assumption would be for v to be distributed log-normally, as the support should be positive. This is equivalent to working directly with $\log(v)$, which is normally distributed. Finding the vector of parameters that maximizes the likelihood is still straightforward and reduces to a single-dimensional optimization problem. Specifically, concentrating the likelihood function allows for a search over α_2 where, for each guess of α_2 , the likelihood-maximizing values of the remaining α parameters are recovered through a least-squares regression and the likelihood-maximizing value of σ^2 is recovered as $\frac{1}{N} \sum_{i=1}^N (\log(v_{i,t}))^2$.

⁵¹ As we are effectively working with $\log(P(Z; \beta))$, parameter estimates are somewhat sensitive to the inclusion of observations with $P(Z; \beta)$ close to zero. For this reason, we trim 2.5% of observations from the tails in estimation. Note that this does not drop any observations used in the welfare calculations, as no observed values of the 1999 price gradient lie in that range.

concave MWTP function (which places a relatively larger value on large increases in crime).

5.3. Welfare analyses using preference-Inversion methods

For the sake of completeness, we now consider a well-known alternative approach to recovering the MWTP function. The preference-inversion method of Bajari and Benkard (2005) offers an alternative framework under which to recover the MWTP function. This framework allows the researcher to altogether avoid estimation in the second stage by inverting first-stage estimates of the hedonic gradient. The opportunity for this inversion comes entirely from assumed restrictions on the functional form of the utility function. Given these assumed restrictions, a single observed consumption decision of Z can be analytically mapped (via inversion) into a unique MWTP function for each household.

A simple version of this inversion is the Horizontal approach discussed in Section 4, where it is assumed that MWTP slope is zero. In this case, each household's observed consumption of Z may be mapped into a unique MWTP intercept without estimation, as the intercept is simply given by the value of the price gradient at the household's point of consumption. To recover the other parameters of the MWTP function, we regress these intercepts on income, race, and market dummies. The results of this Horizontal approach are repeated in Table 7.

A more nuanced version of this inversion approach is the one described in Bajari and Benkard (2005) where a non-linear utility function is specified that yields a constant, but non-zero elasticity of demand, e.g., a Cobb-Douglas utility function that yields a constant elasticity of -1 . Intuitively, if the researcher is willing to assume a constant elasticity of -1 for the MWTP function, then knowing a single point along this function would allow the researcher to recover the entire function. We implement a version of this Unitary-Elasticity approach with our data by specifying utility as:

$$U = \alpha_{0,t} + \alpha_{1,i,t} \log(\zeta - Z_{i,t}) + g(H_{i,t}, \epsilon_{i,t}) + (I_{i,t} - P(Z_{i,t}, H_{i,t}, \epsilon_{i,t}; \beta_t)) \tag{21}$$

In this case, the argument of the utility function is now a measure of safety, $(\zeta - Z_{i,t})$, where ζ is 1.01 times the maximum observed violent crime level. This allows the MWTP function to be downward sloping in

Table 7
WTP for non-marginal changes in violent crime using preference-inversion methods.

| | Buyers with reductions | | | Buyers with increases | | |
|--------------------|------------------------|--------|--------|-----------------------|--------|--------|
| | (n = 24, 791) | | | (n = 12, 900) | | |
| | Average | 25th % | 75th % | Average | 25th % | 75th % |
| Horizontal | 858 | 308 | 1099 | -1194 | -1537 | -282 |
| Unitary elasticity | 828 | 306 | 1075 | -1298 | -1586 | -283 |

These are estimates of willingness to pay (in year-2000 dollars) for the observed 1999 changes in crime. Horizontal estimates use the inversion procedure for a model with a horizontal MWTP function. Unitary Elasticity estimates use the inversion procedure for the model presented in Eq. 21.

the case of a “good” like safety and yields a constant elasticity of -1 with respect to safety in the following MWTP function:⁵²

$$MWTP = \frac{\alpha_{1,i,t}}{(\zeta - Z_{i,t})} \tag{22}$$

By rearranging the first-order condition, one can then perform the following preference inversion to recover $\hat{\alpha}_{1,i,t}$:

$$\hat{\alpha}_{1,i,t} = P'(Z_{i,t}; \beta_1)(\zeta - Z_{i,t}) \tag{23}$$

Results of this Unitary-Elasticity specification may be found in Table 7 and, while the results reflect a downward sloping MWTP function, they are very similar to the Horizontal case.⁵³

6. Conclusion

Researchers regularly ascribe downward-sloping demand curves to households for goods ranging from breakfast cereals to BMWs. In fact, recovering the price elasticity of demand for such goods constitutes one of the main activities undertaken by applied microeconomists. However, because of the difficult endogeneity problems associated with the recovery of the MWTP function using the hedonic technique, the flexibility to estimate downward sloping demand curves has generally not been applied to household demand for local public goods and amenities. Instead, applications of the hedonic method have tended to focus only on the first-stage hedonic price regression; recovering parameters that only yield valid welfare estimates for marginal policies. In order to properly evaluate the welfare effects associated with larger policies, the researcher must recover the structural parameters of the MWTP function. In this paper, we show how easy the empirical implementation of a parametric model can be, following recent advances in the literature. We show that in this framework, the approach is computationally light, easy to implement, and requires no more in terms of data than the standard Rosen estimation approach. In fact, the maximum-likelihood estimator can be simplified to a search over a single parameter.

Using this maximum-likelihood approach and data on violent crime rates in California’s Bay Area, we find that properly accounting for the

shape of the MWTP function has important implications for measuring the welfare effects of non-marginal changes in violent crime. Considering the welfare effects associated with the observed one-year change in crime for those households that purchased a house in 1999, we find that alternative estimation procedures overstate benefits (for those households which experienced a decrease in violent crime) and understate costs (for those households which experienced an increase in violent crime) by over thirty percent. These differences are both statistically and economically significant and consequential for cost-benefit analyses of policies that may have large impacts on future crime rates.

Appendix A. When a closed-form solution for Z exists

In this subsection, we consider a special case of the general model in which Z may be easily isolated in the first-order condition for utility maximization. In this case, the estimation strategy is particularly transparent.

As this simple model specifies a linear gradient, identification requires some cross-market restrictions, as has been well-established in the literature. In this example, we impose that the slope of the MWTP function is constant across markets, while allowing the MWTP intercepts to vary. We additionally define the variance of v to be common across markets. These cross-market restrictions are not generally necessary for identification but rather for this particular linear-quadratic specification.

Continuing with the same utility function specified in (5), we now parameterize the hedonic price gradient as:

$$P'(Z_{i,j}; \beta) = \beta_{1,j} + \beta_{2,j}Z_{i,j} \tag{24}$$

We arrive at the following first-order condition for Z:

$$\alpha_{1,j} + \alpha_2 Z_{i,j} + \alpha_{3,j} X_i + v_{i,j} - \beta_{1,j} - \beta_{2,j} Z_{i,j} = 0 \tag{25}$$

which can be rearranged such that the single endogenous variable, Z, is isolated on the left:

$$Z_{i,j} = \left(\frac{\alpha_{1,j} - \beta_{1,j}}{\beta_{2,j} - \alpha_2} \right) + \left(\frac{\alpha_{3,j}}{\beta_{2,j} - \alpha_2} \right) X_i + \left(\frac{1}{\beta_{2,j} - \alpha_2} \right) v_{i,j} \tag{26}$$

Eq. (26) describes how the consumption of the amenity Z varies with observable household characteristics, X, unobservable preference shocks, v , and parameters of the hedonic price function, β .

Using hats to indicate that $\hat{\beta}$ is known from the first-stage estimation of the hedonic price function and keeping the same distributional assumption that $v \sim N(0, \sigma^2)$, $Z_{i,j}$ is then distributed normally with mean $\left(\left(\frac{\alpha_{1,j} - \hat{\beta}_{1,j}}{\hat{\beta}_{2,j} - \alpha_2} \right) + \left(\frac{\alpha_{3,j}}{\hat{\beta}_{2,j} - \alpha_2} \right) X_i \right)$ and standard deviation $\left(\frac{\sigma}{\hat{\beta}_{2,j} - \alpha_2} \right)$. This reveals a straightforward maximum-likelihood approach for estimating the utility

⁵² Working directly with Z, which is a “bad”, will yield an upward sloping MWTP function. By using a measure of safety, the MWTP function for Z is allowed to be both negative and downward sloping. In Table 7, the Horizontal and unitary-elasticity models yield similar results. We chose the smallest possible value of ζ and note that if larger values of ζ are chosen, the results become even more similar as larger values of ζ dictate flatter MWTP slopes. See Bishop and Timmins (2018) for a similar discussion.

⁵³ The model that is specified in Section 5.2.2 gives rise to a MWTP function with constant elasticity α_2 . One could also implement a preference inversion by imposing a value for this constant elasticity (versus estimating the constant elasticity coefficient as we do in Section 5.2.2). One possible choice would be the assumption $\alpha_2 = 0$. This is equivalent to assuming a zero slope for the MWTP function and yields the Horizontal results shown in Table 7. Another choice would be to impose unitary elasticity and assume that $\alpha_2 = 1$. (As we are working with a “bad” in the fourth quadrant, this elasticity is positive.) However, this arbitrary choice is somewhat hard to justify as the unitary elasticity comes solely from specifying a linear function that must go through the origin. Nonetheless, the analogous results to Table 7 with this restriction are 783, 303, 1033, -1402, -1726, and -288.

Table 8
 Monte Carlo results (common α_1). (“true” parameter values: $\alpha_1 = 3, \alpha_2 = -0.3, \sigma = 0.5$).

| | mean(α_1) | mean(α_2) | mean(σ) | std(α_1) | std(α_2) | std(σ) |
|-------------------------------------|--------------------|--------------------|------------------|-------------------|-------------------|-----------------|
| $j = 2, \gamma_1 = 1, \gamma_2 = 0$ | 3.0035 | -0.3036 | 0.5015 | 0.0709 | 0.0706 | 0.0357 |
| $j = 2, \gamma_1 = 2, \gamma_2 = 0$ | 3.0004 | -0.3006 | 0.5000 | 0.0354 | 0.0347 | 0.0182 |
| $j = 2, \gamma_1 = 3, \gamma_2 = 0$ | 3.0000 | -0.3001 | 0.4998 | 0.0241 | 0.0231 | 0.0127 |
| $j = 2, \gamma_1 = \gamma_2 = 1$ | 3.0015 | -0.3016 | 0.5005 | 0.0460 | 0.0452 | 0.0233 |
| $j = 2, \gamma_1 = \gamma_2 = 2$ | 3.0002 | -0.3003 | 0.4999 | 0.0240 | 0.0222 | 0.0124 |
| $j = 2, \gamma_1 = \gamma_2 = 3$ | 3.0000 | -0.3001 | 0.4997 | 0.0171 | 0.0146 | 0.0091 |
| $j = 5, \gamma_1 = \gamma_2 = 1$ | 3.0002 | -0.3003 | 0.4999 | 0.0342 | 0.0331 | 0.0175 |
| $j = 5, \gamma_1 = \gamma_2 = 2$ | 2.9998 | -0.2999 | 0.4997 | 0.0182 | 0.0159 | 0.0097 |
| $j = 5, \gamma_1 = \gamma_2 = 3$ | 2.9998 | -0.2999 | 0.4997 | 0.0133 | 0.0100 | 0.0074 |
| $j = 10, \gamma_1 = \gamma_2 = 1$ | 3.0002 | -0.3003 | 0.4998 | 0.0309 | 0.0296 | 0.0158 |
| $j = 10, \gamma_1 = \gamma_2 = 2$ | 2.9998 | -0.2999 | 0.4997 | 0.0166 | 0.0141 | 0.0089 |
| $j = 10, \gamma_1 = \gamma_2 = 3$ | 2.9998 | -0.2999 | 0.4997 | 0.0123 | 0.0087 | 0.0069 |
| $j = 50, \gamma_1 = \gamma_2 = 1$ | 3.0000 | -0.3001 | 0.4998 | 0.0285 | 0.0271 | 0.0147 |
| $j = 50, \gamma_1 = \gamma_2 = 2$ | 2.9998 | -0.2999 | 0.4996 | 0.0155 | 0.0128 | 0.0084 |
| $j = 50, \gamma_1 = \gamma_2 = 3$ | 2.9998 | -0.2999 | 0.4996 | 0.0117 | 0.0078 | 0.0066 |

parameters.⁵⁴ This likelihood is given by:⁵⁵

$$\begin{aligned} & \prod_{i=1}^N \ell(\alpha, \sigma; Z_{i,j}, X_{i,j}) \\ & \text{where } \ell(\alpha, \sigma; Z_{i,j}, X_{i,j}) \\ & = \frac{1}{\left(\frac{\sigma}{\hat{\beta}_{2,j} - \alpha_2}\right)\sqrt{2\pi}} \exp \\ & \times \left\{ -\frac{1}{2\left(\frac{\sigma}{\hat{\beta}_{2,j} - \alpha_2}\right)^2} \left(Z_{i,j} - \left(\left(\frac{\alpha_{1,j} - \hat{\beta}_{1,j}}{\hat{\beta}_{2,j} - \alpha_2} \right) + \left(\frac{\alpha_{3,j}}{\hat{\beta}_{2,j} - \alpha_2} \right) X_{i,j} \right) \right)^2 \right\} \end{aligned} \tag{27}$$

Finally, it is worth considering the very special case of the parametric model when estimation is particularly straightforward: the case where the structural parameters may be recovered using least-squares estimation. As an example, consider the specification above with exactly two markets. In this case, Eq. (26) may be estimated using an indirect least squares (ILS) procedure. With the same number of equations as unknown structural parameters,⁵⁶ it becomes a simple matter to recover the structural parameters $\{\alpha_{1,j}, \alpha_2, \alpha_{3,j}, \sigma\}$ from the reduced-form parameters $\{\theta_{0,j}, \theta_{1,j}, \sigma_{u_j}\}$ which are recovered using OLS for each market j . The unique mapping between the two sets is given by:

$$Z_{i,j} = \underbrace{\left(\frac{\alpha_{1,j} - \hat{\beta}_{1,j}}{\hat{\beta}_{2,j} - \alpha_2}\right)}_{\theta_{0,j}} + \underbrace{\left(\frac{\alpha_{3,j}}{\hat{\beta}_{2,j} - \alpha_2}\right)}_{\theta_{1,j}} X_{i,j} + \underbrace{\left(\frac{1}{\hat{\beta}_{2,j} - \alpha_2}\right)}_{u_{i,j}} v_{i,j} \tag{28}$$

where σ_{u_j} is the market-specific standard deviation of $u_{i,j}$.

Appendix B. Monte Carlo evidence

In this Appendix, we provide Monte Carlo evidence on the performance of this estimation approach for the model described in Appendix A. We begin with Monte Carlo simulations of the simplest two-market model. From this starting point, we increase the number of markets and increase the level of heterogeneity in both the market-specific gradient

⁵⁴ In fact, Kahn and Lang (1988) suggest estimating a restricted version of Eq. (26) via non-linear least squares. However, their estimator requires the strong assumption that all of the utility parameters are constant across markets. Additionally, their proposed estimator is only applicable for the subset of cases where a closed-form solution for Z exists and does not generalize to the cases we present in Section 2.1.

⁵⁵ As this model is simply a special case of the model discussed in Section 2.1, the approach of concentrating the likelihood to facilitate a single-parameter search still applies.

⁵⁶ Let L denote the number of elements in X and J denote the number of markets. The reduced-form estimation returns $(J * (L + 1) + J)$ parameters. The number of structural parameters in Eq. (26) is $(J * (L + 1) + 2)$. Therefore, for $J = 1$, this model is underidentified (given the linear price gradient). For $J = 2$, it is exactly identified. For $J \geq 3$, the model is overidentified.

intercepts and slopes. Finally, we allow the MWTP intercept to vary by market.

For the first simulations, the hedonic gradient is given by:

$$P'(Z_{i,j}; \beta) = \beta_{1,j} + \beta_{2,j} Z_{i,j} \tag{29}$$

and the first-order condition for utility maximization is given by:

$$\alpha_1 + \alpha_2 Z_{i,j} + v_{i,j} - P'(Z_{i,j}; \beta) = 0 \tag{30}$$

yielding the MWTP function:

$$P'(Z_{i,j}; \beta) = \alpha_1 + \alpha_2 Z_{i,j} + v_{i,j} \tag{31}$$

and optimal consumption of Z:

$$Z_{i,j} = \left(\frac{\alpha_1 - \beta_{1,j}}{\beta_{2,j} - \alpha_2}\right) + \left(\frac{1}{\beta_{2,j} - \alpha_2}\right)v_{i,j} \tag{32}$$

We allow the number of markets to take on the following values: $j = \{2, 5, 10, 50\}$. We specify that $\beta_{1,j} = 2 + \eta_1$ and $\beta_{2,j} = 0.7 + \eta_2$ where $\eta_1 \sim \gamma_1 * U(-0.3, 0.3)$ and $\eta_2 \sim \gamma_2 * U(-0.15, 0.15)$. γ is allowed to take on the following values: $\gamma_1 = \{1, 2, 3\}$ and $\gamma_2 = \{0, 1, 2, 3\}$.

In all cases, we keep the total number of observations fixed at $n = 5,000$ with observations per market given by $\frac{n}{j}$. The number of Monte Carlo repetitions per experiment is 1,000. We set the structural parameters to the following “true” values: $\alpha_1 = 3, \alpha_2 = -0.3$, and $\sigma = 0.5$.

The results in Table 8 show that there is very little bias in the finite samples, even in the case of only two markets with limited information coming from each market. The standard deviations of the estimated parameters are small relative to the parameters and, more importantly, the efficiency of the estimator is increasing in both market size and level of gradient heterogeneity.

For comparison, we first run the same set of Monte Carlo experiments using the traditional two-step Rosen framework. Results are presented in Table 9. As expected, the estimator performs poorly, particularly when it comes to recovering the slope of the MWTP function, α_2 . In all cases (even with 50 markets and maximum gradient heterogeneity across markets), both the MWTP intercept (α_1) and the standard deviation of the preference shock (σ) are significantly biased downwards. In addition, the MWTP slope is always biased upwards (as expected); in all but two of the experiments, the mean value of the slope takes on a positive value (implying an upward sloping demand curve).

Finally, we run a set of experiments where the MWTP intercept, α_1 , is allowed to vary across markets. We specify that $\alpha_{1,j} \sim U(2, 4)$, while keeping $\alpha_2 = -0.3$ and $\sigma = 0.5$. Note that in this specification, we require heterogeneity in the slope of the gradients across markets and do not estimate the cases where $\gamma_2 = 0$. Our estimator performs well in each case, including the case with only two markets and minimum gradient heterogeneity. The results from these experiments are presented in Table 10.

Table 9
Monte Carlo results (common α_1) - Rosen approach. (“true” parameter values: $\alpha_1 = 3$, $\alpha_2 = -0.3$, $\sigma = 0.5$).

| | mean(α_1) | mean(α_2) | mean(σ) | std(α_1) | std(α_2) | std(σ) |
|-------------------------------------|--------------------|--------------------|------------------|-------------------|-------------------|-----------------|
| $j = 2, \gamma_1 = 1, \gamma_2 = 0$ | 2.0385 | 0.6615 | 0.0980 | 0.0026 | 0.0026 | 0.0003 |
| $j = 2, \gamma_1 = 2, \gamma_2 = 0$ | 2.1381 | 0.5619 | 0.1857 | 0.0044 | 0.0042 | 0.0009 |
| $j = 2, \gamma_1 = 3, \gamma_2 = 0$ | 2.2649 | 0.4350 | 0.2572 | 0.0053 | 0.0049 | 0.0017 |
| $j = 2, \gamma_1 = \gamma_2 = 1$ | 2.0802 | 0.6129 | 0.1455 | 0.0035 | 0.0036 | 0.0007 |
| $j = 2, \gamma_1 = \gamma_2 = 2$ | 2.2557 | 0.4224 | 0.2598 | 0.0055 | 0.0049 | 0.0019 |
| $j = 2, \gamma_1 = \gamma_2 = 3$ | 2.4299 | 0.2332 | 0.3369 | 0.0068 | 0.0050 | 0.0029 |
| $j = 5, \gamma_1 = \gamma_2 = 1$ | 2.1492 | 0.5381 | 0.1984 | 0.0048 | 0.0047 | 0.0012 |
| $j = 5, \gamma_1 = \gamma_2 = 2$ | 2.4131 | 0.2525 | 0.3299 | 0.0068 | 0.0052 | 0.0028 |
| $j = 5, \gamma_1 = \gamma_2 = 3$ | 2.6150 | 0.0358 | 0.4022 | 0.0079 | 0.0047 | 0.0038 |
| $j = 10, \gamma_1 = \gamma_2 = 1$ | 2.1774 | 0.5076 | 0.2163 | 0.0050 | 0.0049 | 0.0014 |
| $j = 10, \gamma_1 = \gamma_2 = 2$ | 2.4654 | 0.1963 | 0.3501 | 0.0071 | 0.0051 | 0.0031 |
| $j = 10, \gamma_1 = \gamma_2 = 3$ | 2.6669 | -0.0187 | 0.4185 | 0.0081 | 0.0045 | 0.0040 |
| $j = 50, \gamma_1 = \gamma_2 = 1$ | 2.2022 | 0.4807 | 0.2309 | 0.0053 | 0.0050 | 0.0015 |
| $j = 50, \gamma_1 = \gamma_2 = 2$ | 2.5067 | 0.1519 | 0.3652 | 0.0073 | 0.0050 | 0.0033 |
| $j = 50, \gamma_1 = \gamma_2 = 3$ | 2.7046 | -0.0582 | 0.4299 | 0.0082 | 0.0043 | 0.0042 |

Table 10
Monte Carlo results (market-specific $\alpha_{1,j}$). (“true” parameter values: $\alpha_2 = -0.3$, $\sigma = 0.5$).

| | mean(α_2) | mean(σ) | std(α_2) | std(σ) |
|-----------------------------------|--------------------|------------------|-------------------|-----------------|
| $j = 2, \gamma_1 = \gamma_2 = 1$ | -0.3531 | 0.5263 | 0.2406 | 0.1209 |
| $j = 2, \gamma_1 = \gamma_2 = 2$ | -0.3139 | 0.5066 | 0.1028 | 0.0524 |
| $j = 2, \gamma_1 = \gamma_2 = 3$ | -0.3068 | 0.5031 | 0.0662 | 0.0345 |
| $j = 5, \gamma_1 = \gamma_2 = 1$ | -0.3277 | 0.5134 | 0.1535 | 0.0775 |
| $j = 5, \gamma_1 = \gamma_2 = 2$ | -0.3084 | 0.5037 | 0.0693 | 0.0359 |
| $j = 5, \gamma_1 = \gamma_2 = 3$ | -0.3045 | 0.5018 | 0.0431 | 0.0233 |
| $j = 10, \gamma_1 = \gamma_2 = 1$ | -0.3221 | 0.5103 | 0.1335 | 0.0675 |
| $j = 10, \gamma_1 = \gamma_2 = 2$ | -0.3068 | 0.5027 | 0.0606 | 0.0316 |
| $j = 10, \gamma_1 = \gamma_2 = 3$ | -0.3036 | 0.5012 | 0.0370 | 0.0204 |
| $j = 50, \gamma_1 = \gamma_2 = 1$ | -0.3193 | 0.5069 | 0.1220 | 0.0616 |
| $j = 50, \gamma_1 = \gamma_2 = 2$ | -0.3061 | 0.5003 | 0.0554 | 0.0290 |
| $j = 50, \gamma_1 = \gamma_2 = 3$ | -0.3033 | 0.4990 | 0.0333 | 0.0186 |

Table 11
WTP for non-marginal changes in violent crime.

| | Buyers with reductions ($n = 24, 791$) | | | Buyers with increases ($n = 12, 900$) | | |
|-----|---|------------|------------|--|------------|------------|
| | Average WTP | 25th % WTP | 75th % WTP | Average WTP | 25th % WTP | 75th % WTP |
| LB | 652 | 294 | 917 | -1710 | -1828 | -290 |
| GMM | 730 | 300 | 992 | -1499 | -1712 | -287 |

These are estimates of willingness to pay (in dollars) for the observed 1999 changes in crime. Welfare estimates are constructed using GMM estimates of the model presented in Eq. 5.

Appendix C. Robinson’s two step estimation details

In order to obtain a consistent estimate of β , we first take the expectation of Eq. (12) with respect to Z ,

$$E[P_{i,t}|Z_{i,t}] = E[H_{i,t}|Z_{i,t}]'\beta_t^H + f_t(Z_{i,t}; \beta_t) + E[\epsilon_{i,t}|Z_{i,t}], \tag{33}$$

and subtract it from Eq. (12). This allows us to write:

$$\tilde{P}_{i,t} = \tilde{H}_{i,t}'\beta_t^H + \epsilon_{i,t} \quad \text{where:} \tag{34}$$

$$\tilde{P}_{i,t} = P_{i,t} - E[P_{i,t}|Z_{i,t}] \quad \text{and} \quad \tilde{H}_{i,t} = H_{i,t} - E[H_{i,t}|Z_{i,t}]$$

Robinson (1988) shows that a consistent estimate of β is obtained when estimating Eq. (34) using Ordinary Least Squares in the second stage of his “two-step” method. For the first stage of the Robinson method, we non-parametrically regress P and H on Z , retain the fitted

values $E[P|Z]$ and $E[H|Z]$, and use them to create \tilde{P} and \tilde{H} . In practice, we specify these expectations as being locally quadratic and estimate them using Weighted Least Squares with weights given by a Gaussian kernel.⁵⁷

Appendix D. Results from the first stage estimation with confidence intervals

Figs. 9, 10, and 11 show the year-specific price functions, price gradients, and second-derivatives of the price functions, respectively. Each plot includes a 99% confidence interval created using 250 bootstrap iterations.

Appendix E. GMM estimation

We also estimate the model described in Section 2.1 via Generalized Method of Moments (GMM). The standard tradeoffs exist between Maximum Likelihood (MLE) and GMM; MLE is more efficient, but GMM does not require a distributional assumption on v . We note that for our model, the computational burden for MLE is considerably lower as, under normality, it is straightforward to concentrate the likelihood and simplify the optimization to a single-parameter search.

To implement GMM, we solve for the preference shock, v , using Eq. 8. Moments can then be constructed based on the discussion in Section 2.2. In practice, we assume that $E[W_1 v] = 0$ and $E[W_2(v^2 - \sigma)] = 0$. W_1 includes the set of market dummies, X , X^2 , X^3 , and interactions among the elements of X , as well as these elements interacted with the market dummies. W_2 includes the set of market dummies. We estimate the parameters, $\{\alpha, \sigma\}$, by minimizing a criterion function consisting of a weighted sum of the empirical analogues of these moments. An efficient weight matrix is chosen using Continuous Updating Efficient GMM as discussed in Hansen et al. (1996).

Welfare estimates for the linear MWTP model using the GMM estimation approach are shown in Table 11 where we also repeat our baseline results for comparison. Overall, the results are reasonably similar to our baseline results for the 25th and 75th percentiles, but somewhat different for more extreme changes in crime as seen in the mean welfare estimates. This reflects the fact that GMM estimates a flatter MWTP function.

⁵⁷ In these estimations, we use the bandwidth of 2.15 times the standard deviation of crime. This is the same bandwidth used in the estimation of Eq. (13) where the regressor is also the rate of violent crime.

References

- Bajari, P., Benkard, C.L., 2005. Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach. *J. Polit. Econ.* 113 (6), 1239–1276.
- Bajari, P., Cooley, J., Kim, K., Timmins, C., 2012. A theory-based approach to hedonic price regressions with time-varying unobserved product attributes: the price of pollution. *Am. Econ. Rev.* 102 (5), 1898–1926.
- Bartik, T.J., 1987. The estimation of demand parameters in hedonic price models. *J. Polit. Econ.* 95 (1), 81–88.
- Bayer, P., McMillan, R., Murphy, A., Timmins, C., 2016. A dynamic model of demand for houses and neighborhoods. *Econometrica* 84 (3), 893–942.
- Bishop, K.C., Timmins, C., 2018. Using panel data to easily estimate hedonic demand functions. *J. Assoc. Environ. Resource Econ.* 5 (3), 517–543.
- Black, S.E., 1999. Do better schools matter? parental valuation of elementary education. *Quart. J. Econ.* 114 (2), 577–599.
- Blumstein, A., Wallman, J., 2006. The crime drop and beyond. *Annual Rev. Law Soc. Sci.* 2, 125–146.
- Bui, L.T.M., Mayer, C.J., 2003. Regulation and capitalization of environmental amenities: evidence from the toxic release inventory in massachusetts. *Rev. Econ. Stat.* 85 (3), 693–708.
- Chay, K.Y., Greenstone, M., 2005. Does air quality matter? evidence from the housing market. *J. Polit. Econ.* 113 (2).
- Chernozhukov, V., Galichon, A., Henry, M., Pass, B., 2017. Single market nonparametric identification of multi-attribute hedonic equilibrium models. Working Paper. MIT.
- Clapp, J., 2003. A semiparametric method for valuing residential location: application to automated valuation. *J. Real Estate Finance Econ.* 27 (3), 303–320.
- Court, A.T., 1939. Hedonic Price Indexes with Automotive Examples. *The Dynamics of Automobile Demand*. New York: General Motors
- Davis, L.W., 2004. The effect of health risk on housing values: evidence from a cancer cluster. *Am. Econ. Rev.* 94 (5), 1693–1704.
- Ekeland, I., Heckman, J.J., Nesheim, L., 2004. Identification and estimation of hedonic models. *J. Polit. Econ.* 112 (1), 109. S60–S
- Ellickson, B., 1971. Jurisdictional fragmentation and residential choice. *Am. Econ. Rev.* 61 (2), 334–339.
- Epple, D., 1987. Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. *J. Polit. Econ.* 95 (1), 59–80.
- Epple, D., Filimon, R., Romer, T., 1984. Equilibrium among local jurisdictions: towards an integrated approach of voting and residential choice. *J. Public Econ.* 24, 281–304.
- Epple, D., Platt, G.J., 1998. Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes. *J. Urban Econ.* 43 (1), 23–51.
- Epple, D., Romano, R.E., 1998. Competition between private and public schools, vouchers, and peer-Group effects. *Am. Econ. Rev.* 88 (1), 33–62.
- Epple, D., Sieg, H., 1999. Estimating equilibrium models of local jurisdictions. *J. Polit. Econ.* 107, 645–681.
- Figlio, D.N., Lucas, M.E., 2004. Whats in a grade? school report cards and the housing market. *Am. Econ. Rev.* 94 (3), 591–604.
- Gamper-Rabindran, S., Timmins, C., 2013. Does cleanup of hazardous waste sites raise housing values? evidence of spatially localized benefits. *J. Environ. Econ. Manage.* 65 (3), 345–360.
- Gayer, T., Hamilton, J.T., Viscusi, W.K., 2000. Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk. *Rev. Econ. Stat.* 82 (3), 439–451.
- Gibbons, S., 2004. The costs of urban property crime. *Econ. J.* 114 (499), 463. F441–F
- Greenstone, M., Gallagher, J., 2008. Does hazardous waste matter? evidence from the housing market and the superfund program. *Q. J. Econ.* 123 (3), 951–1003.
- Griliches, Z., 1961. Hedonic price indexes for automobiles: an econometric of quality change. NBER Chapt. Price Stat. Federal Govern. 173–196.
- Hansen, L.P., Heaton, J., Yaron, A., 1996. Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics* 14 (3), 262–280.
- Heaton, P., 2010. Hidden in plain sight: what the cost of crime research can tell us about investing in police. RAND Center Qual. Occasion. Polic. Paper.
- Heckman, J.J., Matzkin, R.L., Nesheim, L., 2005. Simulation and Estimation of Hedonic Models. In: Kehoe, T., Srinivasan, T.N., Whalley, J. (Eds.), *Frontiers in Applied General Equilibrium*. Cambridge University Press, Cambridge.
- Heckman, J.J., Matzkin, R.L., Nesheim, L., 2010. Nonparametric identification and estimation of nonadditive hedonic models. *Econometrica* 78 (5), 1569–1591.
- Kahn, S., Lang, K., 1988. Efficient estimation of structural hedonic systems. *Int. Econ. Rev. (Philadelphia)* 157–166.
- Kuminoff, N.V., Parmeter, C.F., Pope, J.C., 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *J. Environ. Econ. Strat.* 60 (3), 145–160.
- Lancaster, K.J., 1966. A new approach to consumer theory. *J. Polit. Econ.* 74 (2), 132–157.
- Levitt, S.D., 2004. Understanding why crime fell in the 1990s: four factors that explain the decline and six that do not. *J. Econ. Perspect.* 18 (1), 163–190.
- Linden, L., Rockoff, J.E., 2008. Estimates of the impact of crime risk on property values from Megan's laws. *Am. Econ. Rev.* 98 (3), 1103–1127.
- Nesheim, L., 2015. Identification of multidimensional hedonic models. Working Paper. UCL.
- Palmquist, R.B., 2005. Property Value Models. In: Goran-Maler, K., Vincent, J.R. (Eds.), *The Handbook of Environmental Economics*, Vol. 2. Elsevier.
- Parmeter, C.F., Pope, J.C., 2009. Quasi-experiments and Hedonic Property Value Methods. In: List, J.A., Price, M.K. (Eds.), *Handbook on Experimental Economics and the Environment*. Edward Elgar Publishers.
- Pope, J.C., 2008. Buyer information and the hedonic: the impact of a seller disclosure on the implicit price for airport noise. *J. Urban Econ.* 63 (2), 498–516.
- Poterba, J., 1984. Tax subsidies to owner-occupied housing: an asset-market approach. *Quart. J. Econ.* 99 (4), 729–752.
- Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *J. Polit. Econ.* 82 (1), 34–55.
- Sieg, H., Smith, V.K., Banzhaf, H.S., Walsh, R., 2002. Interjurisdictional housing prices in locational equilibrium. *J. Urban Econ.* 52 (1), 131–153.
- Taylor, L.O., 2003. The Hedonic Method. In: Brown, T., Champ, P., Boyle, K. (Eds.), *A Primer on Nonmarket Valuation of the Environment*. Dordrecht: Kluwer Academic Publishers.
- Yinger, J., 2014. Hedonic markets and sorting equilibria: bid-function envelopes for public services and neighborhood amenities. *J. Urban Econ.* 86, 9–25.