(Under)Mining Local Residential Property Values: A Semiparametric Spatial Quantile Autoregression*

Emir Malikov[1] Yiguo Sun[2] Diane Hite[1]

[1]Auburn University, United States
[2]University of Guelph, Canada

July 9, 2018

Abstract

Rock mining operations, including limestone and gravel production, have considerable adverse effects on residential quality of life due to elevated noise and dust levels resulting from dynamite blasting and increased truck traffic. This paper provides the first estimates of the effects of rock mining—an environmental disamenity—on local residential property values. We focus on the relationship between a house's price and its distance from nearby rock mine. Our analysis studies Delaware County, Ohio which, given its unique features, provides a natural environment for the valuation of property-value-suppressing effects of rock mines on nearby houses. We improve upon the conventional approach to valuating adverse effects of environmental disamenities based on hedonic house price functions. Specifically, in a pursuit of robust estimates, we develop a novel (semiparametric) partially linear spatial quantile autoregressive model which accommodates unspecified nonlinearities, distributional heterogeneity as well as spatial dependence in the data. We derive the consistency and normality limit results for our estimator as well as propose a consistent model specification test. We find statistically and economically significant property-value-suppressing effects of being located near an operational rock mine which gradually decline to insignificant near-zero values at a roughly ten-mile distance. Our estimates suggest that, all else equal, a house located a mile closer to a rock mine is priced, on average, at about 2.3–5.1% discount, with more expensive properties being subject to larger markdowns.

Keywords: Environmental Disamenity, Hedonic Model, Partially Linear, Quantile Regression, Rock Mines, SAR, Semiparametric, Spatial Lag

JEL Classification: C14, C21, R30, Q51

* *Email*: emalikov@auburn.edu (Malikov), yisun@uoguelph.ca (Sun), hitedia@auburn.edu (Hite).

We would like to thank the co-editor, Edward Vytlacil, and three anonymous referees for many insightful comments that helped improve this paper. Any remaining errors are our own.

1 Introduction

This paper provides the first estimates of the effects of rock mining—an environmental disamenity—on local residential property values. Rock mining operations, including limestone rock blasting and gravel mining, have considerable adverse effects on residential quality of life primarily due to elevated noise and dust levels resulting from blasting and increased truck traffic. Exacerbating matters, residential building activity and rock mining are also both pro-cyclical. Further, mining operations naturally seek to minimize their transportation costs by locating closer to their consumers in populated areas (Jaeger, 2006) thus increasing opportunities for opposition from local homeowners and citizen groups due to negative externalities associated with the former.

To valuate the effects of rock mining, we estimate Rosen's (1974) first-stage hedonic house price gradient which has long been used to estimate implicit prices of non-marketable local public goods or, as in our case, public bads from the housing market data. To this end, we focus on the relationship between a house's price and its distance from nearby rock mine. This distance effectively represents environmental amenity/quality, with better quality occurring at farther distances from mines as customarily presumed in hedonic studies. Our analysis focuses on Delaware County, Ohio which, given its unique features, provides a natural environment for the valuation of property-value-suppressing effects (if any) of rock mines on nearby houses. According to the U.S. Census Bureau, Delaware County has been among the two fastest growing counties in the state for the past twenty years. At the same time, given its geology, the county has rich limestone formations that have long been exploited as surface mines.[1] Consequently, residential and commercial expansion in the county has been in conflict with traditional land uses: farming and, especially, rock mining.

In our analysis, we seek to improve upon the conventional approach to valuating adverse effects of environmental disamenities based on hedonic house price functions. Specifically, in a pursuit of robust estimates of property-value effects of rock mines located in the vicinity of residential real estate, we estimate a house valuation function via novel (semiparametric) partially linear spatial quantile autoregressive model. The motivation for developing our model is threefold.

First, our partially linear model allows the distance from a house to nearby rock mine to enter the hedonic house price function in a completely unspecified nonparametric fashion thereby accommodating any potential nonlinearities in the relationship between property values and disamenity. This constitutes a significant improvement over prior studies most of which assume linearity and hence a constant marginal effect of the environmental disamenity on house prices. Few exceptions in earlier work include Harrison & Rubenfeld (1978), Kohlhase (1991), Leggett & Bockstael (2000), Hite et al. (2001), Cohen & Coughlin (2008) and Zabel & Guignet (2012) who model the disamenity quadratically, logarithmically or as a series of range-based dummy variables. In contrast to the latter studies, ours however does not assume the form of nonlinearity *a priori* and instead lets the data determine the nature of functional dependence between the distance to rock mine and house prices. Furthermore, by having the price of a house vary with its distance to mine nonparametrically, one no longer needs to *pre*specify the distance threshold beyond which the disamenity is presumed to have a zero effect on property values. Motivated by the argument that the effects of local disamenities are *local* in nature, the latter is usually done by fixing a spatial radius around a given disamenity thereby defining a circular area to be included in the analysis (e.g., Nelson et al., 1992; Reichert et al., 1992; Hite et al., 2001). In practice, the need to prespecify the radius is oftentimes dictated by the fact that one is more likely to find counterintuitive results if "irrelevant" data from far distances are included in the estimation of a parametric model that inherently cannot accommodate unknown nonlinearities in the property-value effects of disamenities, unless correctly

[1]Source: Ohio Department of Natural Resources.

prespecified. Our model is far more robust to this problem since it assumes no particular form of nonlinearity in the relationship between property values and disamenity.

Second, it is well-known in the real estate literature that environmental disamenities are likely to have heterogeneous impacts on residential property values with larger effects expected in more expensive upscale neighborhoods and more modest effects in less expensive areas (e.g., Reichert et al., 1992; Gayer, 2000). Nonetheless, virtually all earlier attempts at measuring the impact of environmental disamenities on property values have done so by estimating a hedonic house price function at the conditional *mean*. Such an approach delivers the marginal effect on the average house price, which can be rather uninformative from a policy perspective even after controlling for neighborhood characteristics because an "average" may not be representative of actual properties within the same locality, especially in the presence of thick tails of the house price distribution. In order to accommodate heterogeneous effects, we therefore assess the property-value impact of rock mines at different conditional *quantiles* of the house price distribution. We accomplish the latter by estimating a quantile regression model which, besides being more robust to the error distributions including the presence of outliers, allows for *distributional* heterogeneity of the effects of rock mines on property values.

Third, our model explicitly allows for spatial dependence in property values. By estimating a spatially autoregressive hedonic price function, we are able to indirectly control for *unobserved* neighborhood characteristics and shared local amenities (e.g., parks, playgrounds, traffic, air quality, crime, etc.) that affect property values. The spatial lag measuring the average price of neighboring houses serves as a good proxy for these unobserved neighborhood-wide attributes because, owing to their shared nature, they are also priced into the *observed* values of neighboring properties. While these characteristics can be partly controlled for using locality fixed effects, such an approach may be unsatisfactory since it does not let characteristics of neighboring houses affect the price of a given house (Anselin & Lozano-Gracia, 2009). However, by including the spatial lag in a hedonic house pricing function, we are able to accommodate such cross-neighbor effects as can be seen from a reduced form of our model whereby the conditional quantile of house price depends not only on its own attributes but also on its neighbors'. Perhaps more importantly, the spatial lag also contains information about (and thus can proxy for) unobserved *property*-specific attributes such as curb appeal because a given property's value, which is already reflective of its unobserved characteristics, affects its neighboring house's price through the "sales comparison approach" to a real estate appraisal whereby real estate agents base their appraisals of properties on the sale price information for houses in the neighborhood (see the references in Small & Steimetz, 2012). Further, the price information about nearby houses that describes neighbors with similar characteristics can also help partial out the confounding effects of residents' own preferences for environmental quality thereby, at least partly, allowing to alleviate the potential problem of a non-random preference-based self-sorting of property owners into neighborhoods of different environmental quality. Thus, our spatially autoregressive hedonic model is more robust to the omitted variable bias problem, which is a salient problem in the housing-market-based valuations of adverse effects of environmental disamenities and from which the majority of empirical studies suffer.[2] Our spatial-lag-based approach to containing the omitted variable bias in the hedonic price function complements alternative solutions considered in the literature which either (i) require quasi-experimental variation (Chay & Greenstone, 2005; Greenstone & Gallagher, 2008) that is scarce and unavailable in most applications or (ii) are data-intensive requiring the availability of a panel of house sale transactions

[2]Prior papers that have also employed spatial hedonic models are largely limited to Gawande & Jenkins-Smith (2001), Brasington & Hite (2005) and Cohen & Coughlin (2008) although, unlike us, these studies of environmental disamenities focus on more restrictive parametric conditional mean models.

propped up by strong assumptions about the evolution of unobservables (Bajari et al., 2012). Our model therefore offers a practically useful addition to the toolkit of practitioners interested in the hedonic valuation of environmental effects.

Our econometric model itself is a stand-alone contribution to the literature. It constitutes a practically useful fusion of semi/nonparametric quantile methods with models of spatial dependence. While the econometric literature has recently seen a rapid development in the theory of semi/nonparametric estimation of quantile models, most papers however do not admit endogenous explanatory variables (e.g., He & Shi, 1996; Yu & Jones, 1998; He & Liang, 2000; Lee, 2003; Honda, 2004; Kim, 2007) and those that do (e.g., Horowitz & Lee, 2007; Chen & Pouzo, 2009, 2012, 2015; Gagliardini & Scaillet, 2012) usually rule out cross-sectional dependence by focusing on the case of *i.i.d.* data. In this paper, we consider a semiparametric quantile regression in the presence of a special kind of endogeneity induced by the spatial dependence in the outcome variable, i.e., a spatial quantile autoregression, with the errors allowed to be a weakly dependent spatial process with heteroskedasticity that takes a more general form of spatial dependence (including nonlinear) than the widely assumed autoregressive error structure in spatial models (e.g., Kelejian & Prucha, 1998, 1999, 2010). Our model nests several special cases that have been studied in the literature with Su & Yang (2011) and Su & Hoshino (2016) being the two most closely related papers [see Section 2 for more discussion]. Building on Chernozhukov & Hansen (2006), we propose estimating our model via a two-step nonparametric sieve instrumental variable (IV) quantile estimator. Under fairly mild regularity conditions, we show that our estimator is consistent and asymptotically normal. Furthermore, given that our partially linear model nests a more traditional *fully* linear spatial autoregressive model as a special case, one may naturally wish to formally discriminate between the two. To do so, we propose a bootstrap model specification test statistic which provides a vehicle for testing for a fully parametric specification of the spatial autoregression as well as an overall relevancy of some covariates in the model. The motivation for our test statistic comes from Ullah's (1985) nonparametric likelihood-ratio test formulated for a conditional mean model[3] which we extend to the quantile framework along the lines of Koenker & Machado (1999). We show the proposed is a consistent test.

We find statistically and economically significant property-value-suppressing effects of being located near an operational rock mine which gradually decline to insignificant near-zero values at a roughly ten-mile distance. For residential property in the middle of the price distribution, our estimates suggest that, all else equal, a house located a mile closer to a rock mine is predicted to be priced, on average, at about 3.1% discount. The analogous average discounts for houses in the first and third quartiles of price distribution are around 2.3 and 3.4%, respectively. For upscale property in the 0.95th quantile, it is at an astounding 5.1%. As a back-of-the-envelope welfare calculation, the above discount estimates imply the average loss in property value associated with the house being located a mile closer to a rock mine ranging from \$3,691 to \$10,970 for houses within the interquartile range of price distribution. For more expensive neighborhoods in the 0.95th quantile, such losses can be, on average, as high as \$28,410. Applying the estimated statistically significant discounts to house prices at each observation lying within a 10-mile radius from the mine to predict an increase in each property's value if it were moved from its actual location to a (counterfactual) 10-mile distance from the mine, we find the aggregate property value loss associated with rock mining in the area to be \$68.4 million at the median. Overall, using our specification test, we find that the proximity to rock mines *does* matter for residential property values.

The rest of the paper unfolds as follows. We first introduce our econometric model in Section 2, where we outline a two-step estimation methodology for it as well as provide its large-sample

[3]Also see Fan et al. (2001) and Lee & Ullah (2003).

statistical properties. Section 3 presents a model specification test. (We study the finite-sample performance of our proposed estimator and the test statistic in a small set of Monte Carlo simulations in the Supporting Information Appendix.) We discuss the data in Section 4. The empirical results are reported in Section 5. Section 6 concludes.

2 A Partially Linear Spatial Quantile Autoregression

Following Jenish & Prucha (2012) and Qu & Lee (2015), we study spatial processes located on a (possibly) uneven lattice space $D \subseteq R^d$ for some $d \geq 1$.[4] Let $\mathcal{Z}_n = \{(y_{i,n}, \mathbf{x}_{i,n}, \mathbf{z}_{i,n}, u_{i,n}, \varepsilon_{i,n}) : \boldsymbol{l}(i) \in D_n, n \geq 1\}$ be a triangular array of random fields defined on a probability space $(\Omega, \mathcal{F}, P)$ with $D_n \subset D$, where D_n is a finite subset of D, and $\boldsymbol{l}(i)$ refers to the location of the ith spatial unit in D, which is equipped with some distance metric $\varrho(i,j)$. For instance, we can let $\varrho(i,j) = \|\boldsymbol{l}(i) - \boldsymbol{l}(j)\|$ be a Euclidean distance between location $\boldsymbol{l}(i)$ and $\boldsymbol{l}(j)$. Also, let $|U|$ denote the cardinality of a finite subset $U \subset D$. In addition, we denote $\mathcal{F}_{i,n}(s) = \sigma(\varepsilon_{j,n}, \boldsymbol{l}(j) \in D_n, \varrho(i,j) \leq s)$ as the smallest σ-field generated by $\{\varepsilon_{i,n}\}$ located in the s-neighborhood of the spatial unit i. We consider the increasing domain asymptotics as described in the following assumption.

Assumption 1 *The lattice D is infinitely countable with $|D_n| = n$, and $\varrho(i,j) > \varrho_0 > 0$ for any $i \neq j$.*

We consider the following partially linear spatial quantile autoregression (PLSQAR) model for a given quantile index τ:

$$y_{i,n} = \rho_{\tau,0} \sum_{j \neq i} w_{ij,n} y_{j,n} + \mathbf{x}'_{i,n} \boldsymbol{\beta}_{\tau,0} + \alpha_{\tau,0}(\mathbf{z}_{i,n}) + u_{i,n} \quad \forall\, \tau \in (0,1), \tag{2.1}$$

where $y_{i,n}$ is the (scalar) outcome variable of interest; $\mathbf{x}_{i,n}$ and $\mathbf{z}_{i,n}$ are $d_x \times 1$ and $d_z \times 1$ vectors of exogenous covariates, respectively; $\sum_{j \neq i} w_{ij,n} y_{j,n}$ is the endogeneity-inducing spatial lag with $w_{ij,n}$ being the (i,j)-th element of an $n \times n$ non-stochastic spatial weighting matrix $\mathbf{W}_n$ such that $w_{ii,n} = 0$ for all i and $\max_{1 \leq i \leq n} |\lambda_i\{\mathbf{W}_n\}| \leq 1$ where $\lambda_i\{\mathbf{A}\}$ is the ith eigenvalue of some $n \times n$ matrix $\mathbf{A}$; $\rho_{\tau,0} \in (-1,1)$ is a scalar spatial lag parameter; $\boldsymbol{\beta}_{\tau,0}$ is a $d_x \times 1$ vector of constant slope parameters; and $\alpha_{\tau,0}(\cdot)$ is a scalar nonparametric function of $\mathbf{z}_{i,n}$. For identification purposes, $\mathbf{x}_{i,n}$ is assumed to include non-constant regressors only, and hence function $\alpha_{\tau,0}(\cdot)$ subsumes a traditional constant intercept parameter. Therefore, we refer to $\alpha_{\tau,0}(\cdot)$ as the "intercept function." Lastly, $u_{i,n}$ is the quantile error term such that

$$\Pr[u_{i,n} \leq 0 | \mathbf{X}_n, \mathbf{Z}_n, \mathbf{M}_n] = \tau \quad \text{a.s.} \quad \forall\, i = 1, \ldots, n, \tag{2.2}$$

where $\mathbf{X}_n = (\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{n,n})'$ and $\mathbf{Z}_n = (\mathbf{z}_{1,n}, \ldots, \mathbf{z}_{n,n})'$ are $n \times d_x$ and $n \times d_z$ data matrices, respectively; and $\mathbf{M}_n = (\mathbf{m}_{1,n}, \ldots, \mathbf{m}_{n,n})'$ is an $n \times d_m$ instrument matrix with $\mathbf{m}_{i,n}$ being a $d_m \times 1$ vector of valid instruments for the endogenous spatial lag $\sum_{j \neq i} w_{ij,n} y_{j,n}$.

Letting $\mathbf{y}_n = (y_{1,n}, \ldots, y_{n,n})'$ and $\mathbf{u}_n = (u_{1,n}, \ldots, u_{n,n})'$, we can rewrite our model (2.1) in the matrix form as follows

$$\mathbf{y}_n = \rho_{\tau,0} \mathbf{W}_n \mathbf{y}_n + \mathbf{X}_n \boldsymbol{\beta}_{\tau,0} + \boldsymbol{\alpha}_{\tau,0}(\mathbf{Z}_n) + \mathbf{u}_n, \tag{2.3}$$

[4] Some examples of high-dimensional lattices include the case when the unit's spatial location is defined in terms of latitude, longitude and altitude ($d = 3$) or the case when spatial relationships are determined on the basis of unit's state, county, city and the zip code ($d = 4$), etc.

where $\boldsymbol{\alpha}_{\tau,0}(\mathbf{Z}_n) = (\alpha_{\tau,0}(\mathbf{z}_{1,n}), \ldots, \alpha_{\tau,0}(\mathbf{z}_{n,n}))'$. From (2.3), it is evident that, by assuming that the eigenvalues of $\mathbf{W}_n$ do not exceed one in absolute magnitude[5] and that the spatial lag parameter lies within the unit circle, we ensure the non-singularity of $\mathbf{I}_n - \rho_{\tau,0}\mathbf{W}_n$ necessary to guarantee the existence of the reduced form for our model:

$$\mathbf{y}_n = [\mathbf{I}_n - \rho_{\tau,0}\mathbf{W}_n]^{-1} \left(\mathbf{X}_n\boldsymbol{\beta}_{\tau,0} + \boldsymbol{\alpha}_{\tau,0}(\mathbf{Z}_n) + \mathbf{u}_n\right). \tag{2.4}$$

The appeal of our proposed semiparametric PLSQAR model in (2.1) is at least two-fold. First, not only does it accommodate heterogeneity in the spatial relationship by allowing some covariates in the model (namely, $\mathbf{z}_{i,n}$) to affect the outcome variable in a completely unspecified way thereby admitting any potential unit-specific nonlinearities but it also allows for *distributional* heterogeneity of the effects of $\mathbf{X}_n$ and $\mathbf{Z}_n$ on $\mathbf{y}_n$. The latter is accomplished by separate measurements of the spatial relationship at different points of a response distribution. Second, unlike more conventional conditional mean models of spatial dependence, our quantile model is more robust to the error distributions including the presence of outliers. For the estimation of closely related partially linear semiparametric *conditional mean* SAR models recently studied in the literature, e.g., see Su & Jin (2010), Su (2012), Zhang (2013), Sun et al. (2014) and Malikov & Sun (2017). A broader literature on the semiparametric conditional mean estimation in the presence of endogenous regressors includes Li & Stengos (1996), Li & Ullah (1998), Baltagi & Li (2001), Cai et al. (2006), Cai & Li (2008), among others.

Model (2.1) nests several special cases of quantile regressions that have been studied in the literature. Perhaps, the two most closely related models are those by Su & Yang (2011) and Su & Hoshino (2016). Specifically, if nonparametric intercept function $\alpha_{\tau,0}(\cdot)$ does not vary with $\mathbf{z}_{i,n}$ and is constant for any given quantile index τ, i.e., when $\alpha_{\tau,0}(\mathbf{z}_{i,n}) = \alpha_{\tau,0}$ for all $\mathbf{z}_{i,n}$, our model becomes a (more restrictive) *fully* parametric linear spatial quantile autoregression (SQAR) considered by Su & Yang (2011). On the other hand, our model can also be viewed as a special case of Su & Hoshino's (2016) varying-coefficient quantile regression where all parameter functions, except for the intercept, are forced to be constant. However, while their model also features endogenous regressors, it rules out any cross-sectional dependence by focusing on the case of *i.i.d.* data. In contrast, our PLSQAR model relaxes the *i.i.d.* assumption by allowing the spatial dependence in $\mathbf{y}_n$. In the case when the outcome variable exhibits no spatial dependence and hence $\rho_{\tau,0} = 0$, our model is no longer subject to endogeneity and essentially becomes an ordinary partially linear quantile regression which has been rather extensively studied for *i.i.d.* data (e.g., He & Shi, 1996; He & Liang, 2000; Lee, 2003). If one further restricts $\boldsymbol{\beta}_{\tau,0} = \mathbf{0}_{d_x}$, the model collapses to a fully nonparametric quantile regression studied by Yu & Jones (1998). In case of exogenous regressors only, some other closely related models include a varying coefficient quantile regression studied by Honda (2004) and Kim (2007) for *i.i.d.* data and Cai & Xu (2008) for the time-series case.

2.1 Sieve IV Quantile Estimator

Our estimation strategy relies on Chernozhukov & Hansen's (2006) idea whereby the solution to the instrument-based quantile restriction (2.2) is essentially equivalent to the search for $\left(\rho_{\tau,0}, \boldsymbol{\beta}'_{\tau,0}, \alpha_{\tau,0}(\mathbf{z}_{i,n})\right)'$ such that zero is the solution to the usual quantile regression of $y_{i,n} - \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} - \mathbf{x}'_{i,n}\boldsymbol{\beta}_{\tau,0} -$

[5]Which is satisfied if one standardizes a raw spatial weighting matrix by dividing all of its elements by its largest eigenvalue in absolute value.

$\alpha_{\tau,0}(\mathbf{z}_{i,n})$ on exogenous $(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, \mathbf{m}_{i,n})$, i.e.,

$$0 \in \arg\min_{f_\tau \in \mathcal{H}_\tau} \mathbb{E}\left[\zeta_\tau\left(\left(y_{i,n} - \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} - \mathbf{x}'_{i,n}\boldsymbol{\beta}_{\tau,0} - \alpha_{\tau,0}(\mathbf{z}_{i,n})\right) - f_\tau(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, \mathbf{m}_{i,n})\right)\right], \tag{2.5}$$

where $\zeta_\tau(u) \equiv u(\tau - \mathbb{1}\{u < 0\})$ for some $u \in R$ is the so-called "check function" with $\mathbb{1}\{\cdot\}$ being the indicator function, and $f_\tau(\cdot) \in \mathcal{H}_\tau$ is some measurable function.

Chernozhukov & Hansen (2006) pioneered this "instrumental variable quantile regression" approach for a parametric (fully linear) constant-coefficient model. Recently, it has been extended to a broader class of semiparametric varying-coefficient models by Su & Hoshino (2016). Both papers however assume *i.i.d.* data, which is certainly *not* the case in our paper given the spatial dependence in $\mathbf{y}_n$. We show that, under some regularity conditions, the approach nonetheless remains valid even for the spatial data. Different from Su & Yang (2011) who study the fully parametric special case of our model, we do so using the Law of Large Numbers (LLN) and Central Limit Theorem (CLT) for spatial near-epoch dependent (NED) processes derived in Jenish & Prucha (2012). In what follows, we outline the estimation methodology for our PLSQAR model. The asymptotic results along with the necessary assumptions to support them are discussed in Section 2.2.[6]

We approximate unknown nonparametric function using sieves [for an excellent review of the sieve methods, see Chen (2007)]. Specifically, let $\{\phi_1(\cdot), \phi_2(\cdot), \ldots\}$ be a sequence of B-spline series (or the tensor product thereof). Then, for each z, we approximate the unknown intercept function $\alpha_{\tau,0}(z)$ by $\boldsymbol{\phi}_{L_n}(z)'\mathcal{A}_{\tau,0}$ where, for any integer $\kappa > 0$, we denote a $\kappa \times 1$ vector of known basis functions $\boldsymbol{\phi}_\kappa(u) = (\phi_1(u), \ldots, \phi_\kappa(u))'$, and the unknown parameter vector $\mathcal{A}_{\tau,0}$ is of dimension L_n. Hence, we can now rewrite our model in (2.1) as follows

$$y_{i,n} \approx \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} + \mathbf{x}'_{i,n}\boldsymbol{\beta}_{\tau,0} + \boldsymbol{\phi}_{L_n}(\mathbf{z}_{i,n})'\mathcal{A}_{\tau,0} + u_{i,n} \quad \forall\, \tau \in (0,1). \tag{2.6}$$

Motivated by Chernozhukov & Hansen (2006) and Su & Hoshino (2016), we also restrict $\mathcal{H}_\tau$ to the following class of linear functions:

$$\mathcal{H}_\tau = \left\{f_\tau(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, \mathbf{m}_{i,n}) = \mathbf{m}'_{i,n}\boldsymbol{\gamma}_\tau\right\}, \tag{2.7}$$

where $\boldsymbol{\gamma}_\tau$ is a $d_m \times 1$ vector of constant parameters.

The sample counterpart of the objective function in the population instrumental variable quantile regression (2.5) then takes the following form:

$$Q_n(\rho_\tau, \boldsymbol{\beta}_\tau, \mathcal{A}_\tau, \boldsymbol{\gamma}_\tau) \equiv \frac{1}{n}\sum_{i=1}^{n}\zeta_\tau\left(y_{i,n} - \rho_\tau\sum_{j\neq i} w_{ij,n}y_{j,n} - \mathbf{x}'_{i,n}\boldsymbol{\beta}_\tau - \boldsymbol{\phi}_{L_n}(\mathbf{z}_{i,n})'\mathcal{A}_\tau - \mathbf{m}'_{i,n}\boldsymbol{\gamma}_\tau\right). \tag{2.8}$$

Based on the rationale behind (2.5), one is to expect the estimate of $\boldsymbol{\gamma}_\tau$ to be close to zero when the estimate of $\left(\rho_{\tau,0}, \boldsymbol{\beta}'_{\tau,0}, \alpha_{\tau,0}(\cdot)\right)'$ is close to the true population value. Building on this intuition, we can estimate unknown $\left(\rho_{\tau,0}, \boldsymbol{\beta}'_{\tau,0}, \alpha_{\tau,0}(\cdot)\right)'$ in two steps.

[6]Our IV quantile estimation methodology distinctly differs from the increasingly popular *conditional mean* GMM estimators for both fully parametric (Kelejian & Prucha, 1998, 1999, 2010) and semiparametric (Malikov & Sun, 2017; Sun & Malikov, 2018) spatially autoregressive models (with and without spatially autoregressive errors) that make use of not only linear but also quadratic moment conditions. In a quantile regression setup that we consider in this paper, we have not found comparable quadratic moments to use in our IV quantile estimator.

- **Step 1**. For a given value of ρ, we estimate the usual quantile regression of $\dot{y}_{i,n}(\rho) \equiv y_{i,n} - \rho \sum_{j \neq i} w_{ij,n} y_{j,n}$ on exogenous covariates $\mathcal{X}_{i,n} = \left(\mathbf{x}_{i,n}', \mathbf{m}_{i,n}', \boldsymbol{\phi}_{L_n}(\mathbf{z}_{i,n})'\right)'$ to obtain the "profiled" estimates of $\boldsymbol{\theta}_{\tau,0}(\rho) = \left(\boldsymbol{\beta}_{\tau,0}(\rho)', \boldsymbol{\gamma}_{\tau,0}(\rho)', \mathcal{A}_{\tau,0}(\rho)'\right)'$ from

$$\widehat{\boldsymbol{\theta}}_{\tau}(\rho) = \underset{\boldsymbol{\theta}_{\tau}(\rho) \in \boldsymbol{\Theta}}{\arg\min} \ \frac{1}{n} \sum_{i=1}^{n} \zeta_{\tau}\left(\dot{y}_{i,n}(\rho) - \mathcal{X}_{i,n}' \boldsymbol{\theta}_{\tau}(\rho)\right), \tag{2.9}$$

  where we denote $\boldsymbol{\theta}_{\tau}(\rho) = (\boldsymbol{\beta}_{\tau}(\rho)', \boldsymbol{\gamma}_{\tau}(\rho)', \mathcal{A}_{\tau}(\rho)')'$ and $\widehat{\boldsymbol{\theta}}_{\tau}(\rho) = \left(\hat{\boldsymbol{\beta}}_{\tau}(\rho)', \hat{\boldsymbol{\gamma}}_{\tau}(\rho)', \widehat{\mathcal{A}}_{\tau}(\rho)'\right)'$, and $\boldsymbol{\theta}_{\tau,0}(\rho)$ is an interior point of $\boldsymbol{\Theta}$, a compact subset of $R^{d_x + d_m + L_n}$, and is the unique solution to the population counterpart of (2.9):

$$\boldsymbol{\theta}_{\tau,0}(\rho) = \underset{\boldsymbol{\theta}_{0}(\rho) \in \boldsymbol{\Theta}}{\arg\min} \ \mathbb{E}\left[\zeta_{\tau}\left(\dot{y}_{i,n}(\rho) - \mathcal{X}_{i,n}' \boldsymbol{\theta}_{0}(\rho)\right)\right]. \tag{2.10}$$

- **Step 2**. We minimize the weighted norm of $\widehat{\boldsymbol{\gamma}}_{\tau}(\rho)$ estimated in the first step with respect to ρ to obtain our estimator of $\rho_{\tau,0}$:

$$\widehat{\rho}_{\tau} = \underset{\rho \in \Lambda_{\rho}}{\arg\min} \ \widehat{\boldsymbol{\gamma}}_{\tau}(\rho)' \mathbf{V}_n \widehat{\boldsymbol{\gamma}}_{\tau}(\rho), \tag{2.11}$$

  where $\mathbf{V}_n$ is some $d_m \times d_m$ symmetric positive-definite weighting matrix, and $\Lambda_{\rho} \subset (-1, 1)$ is a finite interval consistent with definition of the standardized $\mathbf{W}_n$ as discussed earlier. Correspondingly, the estimators of $\boldsymbol{\beta}_{\tau,0}$ and $\mathcal{A}_{\tau,0}$ are respectively given by

$$\widehat{\boldsymbol{\beta}}_{\tau} = \widehat{\boldsymbol{\beta}}_{\tau}(\widehat{\rho}_{\tau}) \quad \text{and} \quad \widehat{\mathcal{A}}_{\tau} = \widehat{\mathcal{A}}_{\tau}(\widehat{\rho}_{\tau}). \tag{2.12}$$

  Hence, for any given $\mathbf{z}$, the sieve estimator of the unknown intercept function $\alpha_{\tau,0}(\mathbf{z})$ is

$$\widehat{\alpha}_{\tau}(\mathbf{z}) = \boldsymbol{\phi}_{L_n}(\mathbf{z})' \widehat{\mathcal{A}}_{\tau}. \tag{2.13}$$

The implementation of our estimator warrants three remarks. First, assuming that $\mathbf{x}_{i,n}$ and $\mathbf{z}_{i,n}$ are strictly exogenous and relevant, a selection of linearly independent variables from $\mathbf{W}_n\mathbf{X}_n, \mathbf{W}_n\mathbf{Z}_n, \mathbf{W}_n^2\mathbf{X}_n, \mathbf{W}_n^2\mathbf{Z}_n, \ldots$ provides a set of good instruments for the endogenous spatial lag $\mathbf{W}_n\mathbf{y}_n$, because we have $[\mathbf{I}_n - \rho_{\tau,0}\mathbf{W}_n]^{-1} = \sum_{j=0}^{\infty} \rho_{\tau,0}^{j} \mathbf{W}_n^{j}$ in the reduced-form model (2.4). Since we only seek to obtain a consistent nonparametric IV estimator without pursuing optimality, we use $\mathbf{m}_{i,n} = \left[(\mathbf{W}_n\mathbf{X}_n)_i', (\mathbf{W}_n\mathbf{Z}_n)_i'\right]'$ as our instruments, having removed any redundant terms, where $(\mathbf{W}_n\mathbf{A})_i = \sum_{j \neq i} w_{ij,n} a_j$ for $\mathbf{A} = \mathbf{X}_n, \mathbf{Z}_n$. Second, the outlined two-step estimation methodology can be operationalized in the form of a grid search or, alternatively, both steps can be estimated jointly via an automatic numerical search. In either case, it is imperative to impose appropriate box constraints on ρ to ensure that it lies within the unit circle. Third, in the second-step estimation, an obvious practical choice for $\mathbf{V}_n$ is an identity matrix, as suggested by Chernozhukov & Hansen (2006) and Su & Yang (2011). In fact, when $d_m = 1$ and our model is exactly identified, we can show that the limiting distribution of our estimator is expectedly invariant to the choice of $\mathbf{V}_n$. In the case of an over-identified model, one however could improve asymptotic efficiency by weighing $\widehat{\boldsymbol{\gamma}}_{\tau}(\rho)$ using the inverse of its asymptotic covariance matrix, which obviously would first need to be consistently estimated. For tractability purposes, in our paper we set $\mathbf{V}_n = \mathbf{I}_{d_m}$.

2.2 Asymptotic Properties

The derivation of limit results for our proposed estimator requires the following assumptions.

Assumption 2 *(i) $\{(\mathbf{x}_{i,n}, \mathbf{z}_{i,n})\}$ is non-stochastic and uniformly bounded in absolute values; (ii) $u_{i,n} = b_{i,n}(\mathbf{X}_n, \mathbf{Z}_n, \boldsymbol{\varepsilon}_n)$ is a function of $\mathbf{X}_n$, $\mathbf{Z}_n$ and $\boldsymbol{\varepsilon}_n$ such that $\Pr[u_{i,n} \le 0] = \tau$ holds almost surely for all i, and $\boldsymbol{\varepsilon}_n = (\varepsilon_{1,n}, \ldots, \varepsilon_{n,n})$ is an $n \times 1$ vector of random fields with uniformly bounded variances; (iii) $\{u_{i,n}, \boldsymbol{l}(i) \in D_n\}$ is uniformly L_2-NED on $\{\varepsilon_{j,n}, \boldsymbol{l}(i) \in D_n\}$ with the NED coefficients of $\psi(s) = O(s^{-\varsigma})$ for some $\varsigma > d$, and the α-mixing coefficients of $\{\varepsilon_{i,n}\}$ satisfy $\alpha(k,l,r) \le (k+l)^{\upsilon}\,\widehat{\alpha}(r)$ for some $\upsilon \ge 0$ and $\sum_{r=1}^{\infty} r^{d(\upsilon+1)-1}\widehat{\alpha}(r) < \infty$.*

Assumption **2(i)**, also used by Qu & Lee (2015), permits a simple exposition of our assumptions without loss of generality and can be relaxed to allow stochasticity with some bounded moment conditions. Under Assumption **2(ii)–(iii)**, $\{u_{i,n}, \boldsymbol{l}(i) \in D_n\}$ is a weakly dependent spatial process with heteroskedasticity and takes a more general form of spatial dependence (including nonlinear) than the popular spatially autoregressive error structure widely assumed in the literature (e.g., Kelejian & Prucha, 1998, 1999, 2010), which can be easily seen from the proof of Lemma **1** in Appendix A. To conserve space, we refer the reader to Jenish & Prucha (2009, 2012) for definition of the spatial α-mixing and NED process including $\alpha(k,l,r)$ and $\widehat{\alpha}(r)$. Since $\mathbf{X}_n$ and $\mathbf{Z}_n$ are non-stochastic, the stochastic property of $u_{i,n}$ is determined solely by its location $\boldsymbol{l}(i)$ and a nonlinear moving average of $\boldsymbol{\varepsilon}_n$. According to Jenish & Prucha (2012), Assumption **2(iii)** holds if $\max_{1\le i\le n}\mathbb{E}[\varepsilon_{i,n}^2] < M < \infty$ and the overall contributions (i.e., weights) of $\{\varepsilon_{i,n}\}$ in absolute values are ignorable among far-away spatial units. The convergence speeds of the mixing coefficients and the NED coefficients to zero are the same as those in Jenish (2016).

To see the validity of Assumption **2(iii)**, consider an example of $u_{i,n} = \sigma_{i,n}\varepsilon_{i,n}$, where $\{\varepsilon_{i,n}\}$ is an *i.i.d.* error with finite variance and independent of $\{(\mathbf{x}_{i,n}, \mathbf{z}_{i,n})\}$, and $\sigma_{i,n} = \lambda_0 + \lambda_1 \sum_{j\neq i} w_{ij,n} y_{j,n} + \mathbf{x}_{i,n}'\boldsymbol{\lambda}_2 + \lambda_3(\mathbf{z}_{i,n})$. Combining with (2.3)–(2.4), we have

$$\boldsymbol{\sigma}_n = \lambda_0 \mathbf{i}_n + \mathbf{X}_n\boldsymbol{\lambda}_2 + \boldsymbol{\lambda}_3(\mathbf{Z}_n) + \lambda_1\mathbf{G}_n\mathbf{X}_n\boldsymbol{\beta}_{\tau,0} + \lambda_1\mathbf{G}_n\boldsymbol{\alpha}_{\tau,0}(\mathbf{Z}_n) + \lambda_1\mathbf{G}_n\mathcal{E}_n\boldsymbol{\sigma}_n, \tag{2.14}$$

where $\boldsymbol{\sigma}_n = (\sigma_{1,n}, \ldots, \sigma_{n,n})'$, $\mathcal{E}_n = \operatorname{diag}\{\varepsilon_{1,n}, \ldots, \varepsilon_{n,n}\}$, and $\mathbf{i}_n$ is an $n \times 1$ vector of ones. Furthermore, letting $\mathbf{S}_n(\rho) = \mathbf{I}_n - \rho\mathbf{W}_n$ and $\mathbf{G}_n(\rho) = \mathbf{W}_n\mathbf{S}_n(\rho)^{-1}$, we define $\mathbf{S}_n = \mathbf{S}_n(\rho_{\tau,0})$ and $\mathbf{G}_n = \mathbf{G}_n(\rho_{\tau,0})$ the latter of which has a typical element $g_{ij,n}$. If the random matrix $\mathbf{I}_n - \lambda_1\mathbf{G}_n\mathcal{E}_n$ is invertible almost surely,[7] $\sigma_{i,n}$ is an MA(∞) spatial process of $\{\varepsilon_{i,n}\}$. Roughly speaking, $\{\sigma_{i,n}, \boldsymbol{l}(i) \in D_n\}$ is L_2-NED on $\{\varepsilon_{j,n}, \boldsymbol{l}(i) \in D_n\}$ by Proposition 1 in Jenish & Prucha (2012) if $\lim_{s\to\infty}\sup_{n,\boldsymbol{l}(i)\in D_n} \sum_{\boldsymbol{l}(j)\in D_n, \varrho(i,j)>s} |g_{ij,n}| = 0$. Consequently, $\{u_{i,n}, \boldsymbol{l}(i) \in D_n\}$ is L_2-NED on $\{\varepsilon_{j,n}, \boldsymbol{l}(i) \in D_n\}$.

Assumption 3 *(i) $\mathbf{S}_n(\rho)$ is a nonsingular matrix over $\rho \in \Lambda_\rho$, and $\rho_{\tau,0}$ is an interior point of Λ_ρ, a compact subset of R; (ii) there exists a positive integer N such that both $\mathbf{W}_n$ and $\mathbf{S}_n^{-1}(\rho)$ have finite row- and column-sum matrix norms for all $n > N$ and $\rho \in \Lambda_\rho$; (iii) $|w_{ij,n}| \le c_1\varrho(i,j)^{-c_2 d}$ for some positive constants c_1 and $c_2 > \varsigma/d$.*

Assumption **3(i)–(ii)** are the regularity conditions (e.g., Kelejian & Prucha, 2010). Assumption **3(iii)** deviates from Qu & Lee (2015) by assuming gradually decaying spatial weights as the distance

[7] Let $e(\mathbf{A})$ be the largest eigenvalue of $\mathbf{A}$ in the absolute value, where $\mathbf{A}$ is an $n \times n$ matrix with a typical element a_{ij}. Then, $e(\mathbf{A}) \le \|\mathbf{A}\|_1$, where $\|\mathbf{A}\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$ by Seber (2008, Property 4.68). Now, $\|\mathbf{I}_n - \lambda_1\mathbf{G}_n\mathcal{E}_n\|_1 \le |1 - \lambda_1 g_{jj,n}\varepsilon_{j,n}| + |\lambda_1|\max_{1\le j\le n}\sum_{i\neq j}|g_{ij}|\,|\varepsilon_{jn}| < 1$ holds almost surely if $\|\mathbf{G}_n\|_1 < M < \infty$, $\{\varepsilon_{in}\}$ has a compact support, and λ_1 is small enough (Seber, 2008, p.472), where $\|\mathbf{G}_n\|_1 < M < \infty$ is a regularity assumption commonly imposed in the spatial autoregressive literature (e.g., Kelejian & Prucha, 2010).

between two spatial units grows, which includes the case when $|w_{ij,n}| = 0$ if $\varrho(i,j)$ is greater than some threshold value.

Assumption 4 *(i) There exists an $L_n \times 1$ vector $\mathcal{A}_{\tau,0}$ such that*

$$\sup_{\mathbf{z}\in\mathcal{S}_{\mathbf{z}}} \left|\alpha_\tau(\mathbf{z}) - \mathcal{A}'_{\tau,0}\boldsymbol{\phi}_{L_n}(\mathbf{z})\right| \leq M L_n^{-\xi} \tag{2.15}$$

for some $\xi > 2$ as $L_n \to \infty$; (ii)$\{\phi_l(\cdot)\}$ is uniformly bounded over all l such that $\left\|\boldsymbol{\phi}_{L_n}\right\| = \sup_{\mathbf{z}} \sqrt{\sum_{l=1}^{L_n} \phi_l(\mathbf{z})} = O\left(\sqrt{L_n}\right)$.

Since $\mathcal{S}_z$ is a compact set, B-spline tensors can be used to construct the basis functions. Hence, Assumption **4** holds if $\alpha_\tau(\cdot)$ is p-smooth with uniformly bounded derivatives up to order p for some $p > \xi$.

Assumption 5 *(i) The joint probability density function of $\mathbf{u}_n$, denoted by $f_{\mathbf{u}_n}(\mathbf{u})$, is continuously differentiable and all its first-order partial derivatives are uniformly bounded by a positive integrable function $g(\mathbf{u})$ over R^n; (ii) there exists two finite constants $\underline{c}$ and $\bar{c}$ such that $0 < \underline{c} \leq \lambda_{\min}\{\boldsymbol{\Sigma}_\tau(\rho)\} \leq \lambda_{\max}\{\boldsymbol{\Sigma}_\tau(\rho)\} \leq \bar{c} < \infty$ uniformly over $\rho \in \Lambda_\rho$; (iii) $\mathcal{A}_2$ is a nonsingular matrix, where $\boldsymbol{\Sigma}_\tau(\rho)$ and $\mathcal{A}_2$ are respectively defined in* (A.6) *and* (A.9) *in Appendix A.*

Assumption **5(i)** implies that the derivative of the marginal pdf of $u_{i,n}$ is uniformly bounded across index i by the Leibniz rule. We use this assumption in the proof of Lemma **1** in Appendix A to show the existence and uniform boundedness of the pdfs of some intermediate random variables derived from $\mathbf{u}_n$. Assumption **5(ii)** ensures the existence of the estimator calculated in Step 1, while Assumption **5(iii)** ensures the existence of the second-step estimator.

Assumption 6 *As $n \to \infty$, $L_n \to \infty$, $nL_n^{1-2\xi} \to 0$ and $L_n^2/n \to 0$.*

Assumption **6** is an assumption on the smoothing parameter L_n to ensure the consistency of our proposed estimator. Specifically, if we set $L_n = cn^q$ for some $c > 0$, Assumption **6** implies that $0 < 1/(2\xi - 1) < q < 1/2$.

Assumption 7 *$F_{u_{i,n}}(u|\bar{u}_{i,n})$ and $f_{u_{i,n}}(u|\bar{u}_{i,n})$ are, respectively, conditional cdf and pdf of $u_{i,n} = u$ conditional on $\bar{u}_{i,n} = \sum_{j\neq i} g_{ij,n}u_{j,n}$, and $f_{u_{i,n}}(u|\bar{u}_{i,n})$ is uniformly bounded and continuous up to the second-order derivatives with respect to u.*

Assumptions **1–6** are used to show the consistency of our first-step estimator, whereas Assumption **7** is used to derive the asymptotic normality results of the second-step estimator. We state our theorems below and relegate their proofs to Appendix A.

Theorem 1 *Under Assumptions **1–6**, we have* $\max_{\rho\in\Lambda_\rho} \left\|\widehat{\boldsymbol{\theta}}_\tau(\rho) - \boldsymbol{\theta}_{\tau,0}(\rho)\right\| = O_p\left(\sqrt{L_n/n}\right)$.

Theorem **1** states that the first-step estimator is uniformly consistent with respect to $\rho \in \Lambda_\rho$, where $\|\cdot\|$ refers to the Euclidean norm. Now we are in a position to derive the limit result for the second-step estimator.

Theorem 2 *Under Assumptions **1–7**, we have*

$$\sqrt{n}\boldsymbol{\Sigma}_n^{-1/2}\begin{pmatrix} \widehat{\rho}_\tau - \rho_{\tau,0} \\ \widehat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_{\tau,0} \\ \widehat{\gamma}_\tau \end{pmatrix} \xrightarrow{d} \mathbb{N}\left(\mathbf{0}, \mathbf{I}_{1+d_x+d_m}\right),$$

and at a give point $\mathbf{z}$

$$\sqrt{n/\omega_{n,\tau}}\left(\widehat{\alpha}_\tau\left(\mathbf{z}\right) - \alpha_{\tau,0}\left(\mathbf{z}\right)\right) \xrightarrow{d} \mathbb{N}\left(0,1\right),$$

where $\boldsymbol{\Sigma}_n$ *and* $\omega_{n,\tau}$ *are defined in the proof of this theorem in Appendix A.*

From the proof of this theorem, we see that $\boldsymbol{\Sigma}_n$ is a nonsingular matrix under Assumption **5(ii)–(iii)** and that $\omega_{n,\tau} = O\left(\sqrt{L_n}\right)$.

Remark 1. We study the finite-sample performance of our proposed two-step estimator in a small set of Monte Carlo simulations, the discussion of which is relegated to the Supporting Information Appendix. Overall, the results are encouraging, and simulation experiments support our asymptotic results.

3 Specification Testing

We next consider a model specification test which permits testing several useful hypotheses. Specifically, for a τth spatial quantile autoregression model written as

$$y_{i,n} = q\left(\sum_{j\neq i} w_{ij,n}y_{j,n}, \mathbf{x}_{i,n}, \mathbf{z}_{i,n}, \tau\right) + u_{i,n} \equiv q_i(\tau) + u_{i,n}, \tag{3.1}$$

we consider the following null hypotheses about the form of $q_i(\tau)$:

$$\mathrm{H}_0(\mathrm{i}): \quad q_i(\tau) = \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} + \mathbf{x}_{i,n}'\boldsymbol{\beta}_{\tau,0} + (1, \mathbf{z}_{i,n}')'\boldsymbol{\delta}_{\tau,0} \tag{3.2}$$

$$\mathrm{H}_0(\mathrm{ii}): \quad q_i(\tau) = \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} + \mathbf{x}_{i,n}'\boldsymbol{\beta}_{\tau,0} + \delta_{\tau,0}, \tag{3.3}$$

against the alternative hypothesis (i.e., the PLSQAR model):

$$\mathrm{H}_1: \quad q_i(\tau) = \rho_{\tau,0}\sum_{j\neq i} w_{ij,n}y_{j,n} + \mathbf{x}_{i,n}'\boldsymbol{\beta}_{\tau,0} + \alpha_{\tau,0}(\mathbf{z}_{i,n}). \tag{3.4}$$

Alternatively, the above null and alternative hypotheses can be rewritten as follows: $\mathrm{H}_0(\mathrm{i})$: $\Pr\left[\alpha_{\tau,0}(\mathbf{z}_{i,n}) = (1, \mathbf{z}_{i,n}')'\boldsymbol{\delta}_{\tau,0}\right] = 1$ for some $\boldsymbol{\delta}_{\tau,0} \in R^{1+d_z}$ against H_1 : $\Pr\left[\alpha_{\tau,0}(\mathbf{z}_{i,n}) = (1, \mathbf{z}_{i,n}')'\boldsymbol{\delta}_{\tau}\right] < 1$ for any $\boldsymbol{\delta}_\tau \in R^{1+d_z}$, and $\mathrm{H}_0(\mathrm{ii})$: $\Pr\left[\alpha_{\tau,0}(\mathbf{z}_{i,n}) = \delta_{\tau,0}\right] = 1$ for some $\delta_{\tau,0} \in R$ against H_1 : $\Pr\left[\alpha_{\tau,0}(\mathbf{z}_{i,n}) = \delta_\tau\right] < 1$ for any $\delta_\tau \in R$. The null in (3.2) is meant to test for linearity of the conditional quantile function in $\mathbf{z}_{i,n}$. In practice, one may choose any desired *parametric* specification for the intercept function $\alpha_{\tau,0}(\cdot)$ to test against the nonparametric alternative in (3.4). The second null in (3.3) is essentially the test of overall relevancy of $\mathbf{z}_{i,n}$.

To test these hypotheses, we essentially propose a nonparametric likelihood-ratio test based on the comparison of the restricted and unrestricted models. The motivation for our test statistic comes from Ullah's (1985) nonparametric test that compares residual sums of squares under the null and the alternative (also see Fan et al., 2001; Lee & Ullah, 2003). The idea behind this test, which is formulated for a conditional mean model, can be extended to the conditional quantile framework along the lines of Koenker & Machado (1999) whereby the estimated residual sum of check functions effectively plays the role of the residual sum of squares. Specifically, for any given quantile index τ, we consider the following residual-based test statistic:

$$T_n = \frac{RSC_{0,\tau} - RSC_{1,\tau}}{RSC_{1,\tau}}, \tag{3.5}$$

where $RSC_{0,\tau}$ is the residual sum of check functions under H_0 computed as $RSC_{0,\tau} = \sum_{i=1}^{n} \zeta_\tau(\widetilde{u}_{i,n})$ with $\widetilde{u}_{i,n} = y_{i,n} - \widetilde{q}_i(\tau)$ being the quantile residual defined as the difference between $y_{i,n}$ and the consistent estimate of $q_i(\tau)$ under either of the two null hypotheses in (3.2)–(3.3); and $RSC_{1,\tau}$ is the residual sum of check functions under H_1 computed as $RSC_{1,\tau} = \sum_{i=1}^{n} \zeta_\tau(\widehat{u}_{i,n})$, where $\widehat{u}_{i,n}$ is the residual from our second-step estimator, i.e., $\widehat{u}_{i,n} = y_{i,n} - \widehat{q}_i(\tau) = y_i - \widehat{\rho}_\tau \sum_{j\neq i} w_{ij,n} y_{j,n} - \mathbf{x}'_{i,n}\widehat{\boldsymbol{\beta}}_\tau$ $-\widehat{\alpha}_\tau(\mathbf{z}_{i,n})$. The estimated residuals under H_0 can be obtained via Su & Yang's (2011) estimator.

Theorem 3 *Under Assumptions **1–5**, under H_0 we have*

$$nT_n = \frac{\mathcal{B}'_{n,1}\boldsymbol{\Sigma}_\tau^{-1}\mathcal{B}_{n,1} - \mathcal{B}'_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}}{2n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\zeta_\tau\left(u_{i,n}\right)\right]} + o_p\left(1\right) = O_p\left(L_n\right),$$

where $\mathcal{B}_{n,0} = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\varphi_\tau\left(u_{i,n}\right)$ with $\boldsymbol{\chi}_{i,n} = (\mathbf{x}'_{i,n}, \mathbf{z}'_{i,n}, \mathbf{m}'_{i,n})'$ and $\varphi_\tau(u) = \tau - \mathbb{1}\left\{u < 0\right\}$, $\boldsymbol{\Sigma}_{\tau,0} = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} f_{u_{i,n}}\left(0\right)\boldsymbol{\chi}_i\boldsymbol{\chi}'_i$, $\mathcal{B}_{n,1} = n^{-1/2}\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_\tau\left(u_{i,n}\right)$, $\boldsymbol{\Sigma}_\tau = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f_{u_{i,n}}\left(0\right)\mathcal{X}_{i,n}\mathcal{X}'_{i,n}$; and under H_1 we have $T_n \xrightarrow{p} \left(\varpi_0 - \varpi_1\right)/\varpi_1 > 0$, where ϖ_0 and ϖ_1 are finite constants defined in eq. (A.16) in Appendix A.

The proof of this theorem derived in Appendix A shows that $T_n = O_p\left(L_n/n\right)$ under H_0 and that T_n converges to a non-zero positive constant under H_1. Thus, T_n is a consistent test. Intuitively, the test statistic is expected to converge to zero under the null hypothesis and a positive constant under the alternative hypothesis. Hence, the test is one-sided.

For any given L_n, we see that $\mathcal{B}_{n,1} \xrightarrow{d} \mathbb{N}\left(\mathbf{0}, \boldsymbol{\Omega}_\tau\right)$, where $\boldsymbol{\Omega}_\tau = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(\mathcal{X}_{i,n}\mathcal{X}'_{j,n}\times \left[\mathbb{E}\left(\mathbb{1}\left(u_{i,n} < 0\right)\mathbb{1}\left(u_{j,n} < 0\right)\right) - \tau^2\right]\big) = O\left(1\right)$ under Assumption **2(iii)**, and $\boldsymbol{\Omega}_\tau = \tau\left(1-\tau\right)\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\mathcal{X}_{i,n}\mathcal{X}'_{i,n}$ when $\{u_{i,n}\}$ is independently distributed.

Specifically, when $\{u_{i,n}\}$ is an *i.i.d.* sequence, we have under H_0

$$\frac{f_u\left(0\right)}{\tau\left(1-\tau\right)}\mathcal{B}'_{n,1}\boldsymbol{\Sigma}_\tau^{-1}\mathcal{B}_{n,1} \xrightarrow{d} \chi^2\left(d_x + d_m + L_n\right)$$
$$\frac{f_u\left(0\right)}{\tau\left(1-\tau\right)}\mathcal{B}'_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0} \xrightarrow{d} \chi^2\left(d_x + d_m + d_z\right)$$

and, because $\boldsymbol{\chi}_{i,n}$ is a part of $\mathcal{X}_{i,n}$, we obtain

$$\frac{f_u\left(0\right)}{\tau\left(1-\tau\right)}\left(\mathcal{B}'_{n,1}\boldsymbol{\Sigma}_\tau^{-1}\mathcal{B}_{n,1} - \mathcal{B}'_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}\right) \xrightarrow{d} \chi^2\left(L_n\right),$$

from where we have that

$$\frac{f_u\left(0\right)\mathbb{E}\left[\zeta_\tau\left(u_{i,n}\right)\right]}{2\tau\left(1-\tau\right)}nT_n \xrightarrow{d} \chi^2\left(L_n\right) \tag{3.6}$$

for a given L_n. With $\{\widetilde{u}_{i,n}\}$ denoting the residuals calculated under H_0, we can then construct a modified test statistic based on (3.6) as

$$J_n^* = \frac{\widehat{f}_u\,(0)\, n^{-1} \sum_{i=1}^{n} \zeta_\tau\,(\widetilde{u}_{i,n})}{2\tau\,(1-\tau)} nT_n,$$

where $\widehat{f}_u\,(0)$ is the kernel estimator of $f_u\,(0)$. We reject H_0 at the significance level a if $J_n^* \geq c_a$, where $\Pr\left[\chi^2\,(L_n) \geq c_a\right] = a$.

When $\{u_{i,n}\}$ is independent but *not* identically distributed, $\boldsymbol{\Omega}_\tau \neq \boldsymbol{\Sigma}_\tau$, and $\mathcal{B}_{n,1}' \boldsymbol{\Sigma}_\tau^{-1} \mathcal{B}_{n,1}$ can be rewritten as a linear combination of $d_x + d_m + L_n$ independent chi-squared random variables but does not converge to the standard χ^2 distribution with $(d_x + d_m + L_n)$ degree of freedoms. For this case, we cannot derive the exact limit distribution for our test statistic under H_0. Therefore, in this case, we suggest using bootstrap to approximate the null distribution of J_n^*, especially given that residual-based nonparametric tests are well-known to perform rather poorly in finite samples when relying on asymptotic critical values. Bootstrap methods however offer a means to improve their finite-sample performance. For a fixed $\tau \in (0,1)$, we use the following wild (residual) bootstrap procedure adjusted to suit the asymmetric loss function used in the quantile estimation:[8]

(1) Estimate the restricted model under either of the two nulls in (3.2)–(3.3) to obtain residuals $\{\widetilde{u}_{i,n};\ i = 1, \ldots, n\}$. Compute the modified test statistic as $J_n = \left(n^{-1} \sum_{i=1}^{n} \zeta_\tau\,(\widetilde{u}_{i,n})\right) nT_n$.

(2) Generate two-point wild bootstrap errors by setting $u_{i,n}^b = \omega_1 \times |\widetilde{u}_{i,n}|$ with probability $(1-\tau)$ and $u_{i,n}^b = \omega_2 \times |\widetilde{u}_{i,n}|$ with probability τ, where $\omega_1 = 2(1-\tau)$ and $\omega_2 = -2\tau$.

(3) Construct the bootstrap sample $\{y_{i,n}^b, \sum_{j\neq i} w_{ij,n} y_{j,n}^b, \mathbf{x}_{i,n}, \mathbf{z}_{i,n};\ i = 1, \ldots, n\}$, where $y_{i,n}^b$ is generated from the restricted model under the appropriate null:

$$\mathbf{y}_n^b = \begin{cases} [\mathbf{I}_n - \widetilde{\rho}_\tau \mathbf{W}_n]^{-1} \left(\mathbf{X}_n \widetilde{\boldsymbol{\beta}}_\tau + [\mathbf{i}_n, \mathbf{Z}_n] \widetilde{\boldsymbol{\delta}}_\tau + \mathbf{u}_n^b\right) & \text{for } \mathrm{H}_0(\mathrm{i}) \\ [\mathbf{I}_n - \widetilde{\rho}_\tau \mathbf{W}_n]^{-1} \left(\mathbf{X}_n \widetilde{\boldsymbol{\beta}}_\tau + \mathbf{i}_n \widetilde{\delta}_\tau + \mathbf{u}_n^b\right) & \text{for } \mathrm{H}_0(\mathrm{ii}), \end{cases}$$

where $\mathbf{y}_n^b = (y_{1,n}^b, \ldots, y_{n,n}^b)'$ and $\mathbf{u}_n^b = (u_{1,n}^b, \ldots, u_{n,n}^b)'$.

(4) Reestimate both the restricted and unrestricted models using the bootstrap sample from step (3) to obtain bootstrap residuals $\{\widetilde{u}_{i,n}^b;\ i = 1, \ldots, n\}$ and $\{\widehat{u}_{i,n}^b;\ i = 1, \ldots, n\}$ under H_0 and H_1, respectively.

(5) Compute the bootstrap modified test statistic $J_n^b = 2\tau\,(1-\tau)\left(n^{-1} \sum_{i=1}^{n} |\widetilde{u}_{i,n}|\right) nT_n^b$ with $T_n^b = \left(RSC_{0,\tau}^b - RSC_{1,\tau}^b\right)/RSC_{1,\tau}^b$, where $RSC_{0,\tau}^b = \sum_{i=1}^{n} \zeta_\tau(\widetilde{u}_{i,n}^b)$ and $RSC_{1,\tau}^b = \sum_{i=1}^{n} \zeta_\tau(\widehat{u}_{i,n}^b)$.

(6) Repeat steps (2)–(5) B times.

(7) Use the empirical distribution of $B+1$ bootstrap statistics $\{J_n^b\}$, where the first bootstrap test statistic equals the test statistic calculated from the raw data in Step 1, to obtain the upper $a \times 100$th percentile value c_a for a given $a \in (0,1)$. Use this c_a to approximate the upper percentile (critical) value of the modified test statistic J_n under H_0. We reject H_0 if the test statistic is greater than c_a.

[8] Feng et al. (2011) show that a traditional wild bootstrap procedure is invalid for quantile estimators due to nonlinear score functions associated with the check-function-based objective function. Alternatively, Sun (2006) introduces a modified wild bootstrap method applicable to testing in the quantile regression framework.

Theorem 4 *Under Assumptions **1–5**, (i) $J_n^b = O_p(L_n/n)$ under both H_0 and H_1; (ii) under H_0 we have*

$$d_2\left(F^b\left(J_n^b\right), F(J_n)\right) \xrightarrow{p} 0,$$

where $F^b(\cdot)$ is the conditional asymptotic distribution of the bootstrap test J_n^b given $\{(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, y_{i,n})\}$, $F(\cdot)$ is the asymptotic distribution of J_n, and $d_2(G_1, G_2)$, the Mallows distance, is the infimum of $\mathbb{E}(X-Y)^2$ over all joint distributions of the random variable (X, Y) whose marginal distribution functions are G_1 and G_2, respectively; (iii) under H_1, we have $\Pr[J_n > c_a] \to 1$.

Theorem **4** states that our bootstrap procedure provides a consistent test because J_n^b and J_n have the same asymptotic distribution under H_0 and $\Pr[J_n > c_a] \to 1$ under H_1.

Remark 2. Note that, in the case when $\tau = 0.5$, we have $\sum_{i=1}^n \zeta_\tau(\widetilde{u}_{i,n}) = 2\tau(1-\tau)\sum_{i=1}^n |\widetilde{u}_{i,n}|$ which implies that the bootstrap results based on T_n and T_n^b are the same as those based on J_n and J_n^b in our bootstrap procedure. Hence, for *median* regression models, we can conduct the bootstrap specification testing directly based on the T_n statistic in place of its modified analogue J_n.

Monte Carlo simulations (discussed in the Supporting Information Appendix) show that the bootstrap test based on J_n has quite an accurate size and exhibits superb power which rises with the sample size, as expected. Lastly, note that the above bootstrap method does not apply to the general case when $\{u_{i,n}\}$ exhibits cross-sectional weak dependence. The bootstrap procedure for such a case is beyond the scope of this paper, and we leave it for future research.

4 Data

Our data come from Delaware County Auditor's Office and were obtained in the form of ArcGIS parcel shapefiles. Each parcel record contains information about house and other property characteristics such as house and lot size, number of rooms, etc. (see Table 1 for a full self-descriptive list of variables). Based on land-use codes, we retain only records containing arm's length single-family home transactions. We do so because hedonic models require competitive housing markets with buyers and sellers whose willingnesses to pay and accept are formed based on property characteristics only. Our operational sample includes 5,500 sale transactions that took place in the county during the 2009:1–2011:3 period (roughly, two years).

There are four rock mines in the county, three of which are no longer operational. All are surface mines. They were located from geographic coordinates of parcels owned by the mining companies (Ohio Department of Natural Resources, 2010, 2011) and were further verified using Google Earth. The only operational mine (state mine number: Del-5) also happens to be the largest of all by an order of magnitude. It is located in the Southwestern part of the county near the city of Delaware and is about 510 acres large,[9] which is almost triple the size of an average farm in the county (187 acres). In the case of Delaware County, all mines are limestone (but colloquially called gravel mines) and thus are subject to dynamite blasting which creates a far greater nuisance than other types of mines such as composite mines. Given that other mines in the county were no longer in operation by the period of our study and hence did not generate noise, dust and traffic, in our analysis we solely focus on the operational Del-5 mine, which is not only very large but is also located in an area of high urban growth.

[9]Based on Google Earth Pro measurements.

Table 1. Data Summary Statistics

Variable	Units	Mean	5th Perc.	Median	95th Perc.
House Price	thousands $	258.42	64.00	232.49	552.50
Distance to Rock Mine	thousands ft.	49.12	12.92	51.14	80.27
Square Footage	ft.^{2}	2,452.99	1,188.00	2,360.00	4,054.05
Acreage	acres	0.78	0.15	0.30	3.18
Age	years	20.42	0	10	108
Story Height	cardinal number	1.79	1	2	2
# of Bedrooms	cardinal number	3.58	3	4	4
# of Bathrooms	cardinal number	2.95	1	3	5
# of Fireplaces	cardinal number	0.83	0	1	1
Garage Capacity	cardinal number	1.29	0	2	3
Attached Garage	binary indicator	0.551			
Full Basement	binary indicator	0.447			
Partial Basement	binary indicator	0.457			
Attic	binary indicator	0.095			
Central A/C	binary indicator	0.885			
Black Population Share	% pt.	3.27	0.00	1.88	11.11
Median Income	thousands $	80.04	36.40	81.20	113.00
Property Tax Rate	% pt.	1.87	1.39	1.92	2.23

The last three variables are at the Census block group level.

Because our data are explicitly georeferenced, we use a standard software routine to calculate straight-line distances from each property to the mine centroid. This distance proxies environmental amenity associated with rock mining, with better quality occurring at farther distances from mines. We opt for such a measure over the alternative measures of environmental quality associated with disamenities such as the number of disamenities within a certain distance of a property because, in our case, we have a single occurrence of a large disamenity spread widely throughout the area. Further, since our econometric model allows environmental impacts to be nonlinear, the use of straight-line distances as a measure of environmental quality does not appear that problematic.

We also match our data with the neighborhood-specific demographic variables at the Census block level from the U.S. Census Bureau. Specifically, we include the black[10] population share, median income and the property tax rate in the neighborhood. We use these variables as observable controls for neighborhood characteristics (in addition to the spatial lag term as discussed in the introduction). We opt for these continuous measures of neighborhood characteristics over discrete locality fixed effects primarily out of computational considerations because quantile estimation is known to performs rather poorly in the presence of multiple binary covariates.

5 Empirical Results

We estimate the hedonic house valuation function in the form of our PLSQAR model in (2.1), where we let the distance to nearby rock mine enter the function nonparametrically as a "z" variable with the rest of hedonic attributes included parametrically as "x" variables. All right-hand-side covariates appear in levels except for square footage and acreage to which we apply the logarithmic transformation. In the case of the number of bedrooms, bathrooms and age, we

[10] Variables for other non-white population groups have been consistently found to be insignificant, and their exclusion has affected the results in no material way.

also include quadratic terms. Following the literature, we take the logarithm of the left-hand-side house price (the "y" variable) thereby facilitating the interpretation of marginal effects in terms of percentages, allowing for nonlinearities and ensuring the outcome variable can take any real value.

Given the highly uneven distribution of houses in space, we use a distance-based k-nearest-neighbor type of spatial weighting matrices to model spatial relationship across properties. The latter helps ensure that each house gets neighbors whose prices are deemed "relevant" (by getting relatively large weights) in predicting its value. The use of alternative distance-based weighting matrices, where the spatial weights are decaying functions of distance, leads to an undesirable situation when houses in highly urbanized localities have multiple "relevant neighbors that are assigned large weights and houses in a sparsely populated countryside hardly have any such "relevant" neighbors, which obviously is inaccurate because appraisers are willing to look far for comparable properties when valuating houses in rural areas. We select the number of nearest neighbors that minimizes the AIC criterion for the median model. The data favor $k = 5$, which we use throughout.

When estimating the model, we approximate the unknown nonparametric intercept function $\alpha_{\tau,0}(\cdot)$ via cubic B-spline sieves, the order of approximation for which (in this case, the number of knots) is also selected by minimizing AIC. Throughout, we use spatial lags of continuous house-specific attributes (log square footage and log acreage) as our instruments. We do not include lags of other exogenous attributes into the instrument set because they are discrete and lead to severe multicollinearity and convergence problems.

Since the objective of our paper is to assess property-value-suppressing effects of rock mines on nearby property (and in order to conserve space), in what follows we primarily focus on the results concerning the relationship between a house's price and its distance from the mine. Consistent with the notion that rock mines are an environmental disamenity that creates negative externalities such as dust, noise and additional traffic, our expectation is the *positive* relationship between the two variables implying that the houses located farther from mines would be appraised at higher values. (The results pertaining to other house attributes are relegated to the Supporting Information Appendix.)

As discussed earlier, most studies pursuing the housing-market-based valuation of adverse welfare effects of environmental disamenities estimate a linear hedonic price function, which rather restrictively assumes constant marginal impact of the disamenity on house prices. Few papers that do explore potential nonlinearities have largely favored quadratic or logarithmic forms (e.g., Kohlhase, 1991; Leggett & Bockstael, 2000; Hite et al., 2001) which, given their reliance on an *a priori* functional form assumption, are still subject to potential misspecification. We circumvent these problems by letting the distance between the house and a rock mine (z) enter the house valuation function in a nonparametric fashion [through an unspecified intercept function $\alpha_{\tau,0}(\cdot)$] thereby accommodating any potential nonlinearities in the relationship between (log) property values and the (level of) distance to the mine. We first examine the sensitivity of empirical results to potential functional-form misspecification of $\alpha_{\tau,0}(\cdot)$. To do so, in addition to our semiparametric PLSQAR model of house prices, we also estimate a *fully* parametric SQAR model under the following three specifications of the intercept function: (i) $\alpha_{\tau,0}(z) = a_{0,\tau} + a_{1,\tau}z + a_{2,\tau}z^2$, (ii) $\alpha_{\tau,0}(z) = a_{0,\tau} + a_{1,\tau}\log(z)$ and (iii) $\alpha_{\tau,0}(z) = a_{0,\tau} + a_{1,\tau}z$. These specifications imply quadratic, logarithmic and linear functional forms of the relationship between the log price and (the level of) z, respectively. Comparing the results from our flexible PLSQAR model, which lets the data determine the shape of $\alpha_{\tau,0}(\cdot)$, to those from a parametric model under these three specifications enables us to empirically assess the extent to which the hedonic estimates of property-value-suppressing effects of rock mines on nearby houses are sensitive to "correct" functional form specification of the house price function. Such a comparison is especially interesting given the wide popularity

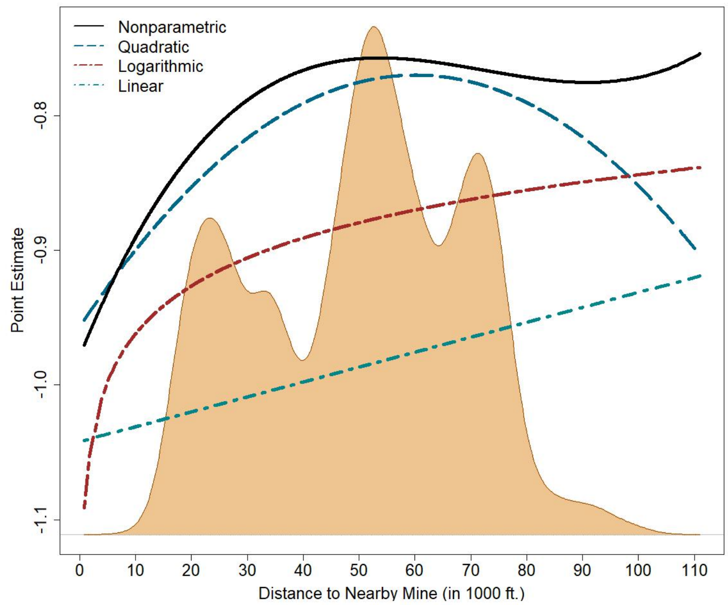

Figure 1. Estimated Intercept Functions of the Distance to Rock Mine for the Conditional Median Model [*Note:* Shaded is the kernel density of the distance variable]

of linear, quadratic and logarithmic parameterizations in the literature. The parametric model under all three specifications of $\alpha_{\tau,0}(\cdot)$ is estimated via a two-step procedure following Su & Yang (2011). To conserve space, we focus on the median quantile ($\tau = 0.50$) of the log house price when comparing these alternative models.

Figure 1 plots the estimated intercept function across the four models.[11] Our preferred PLSQAR model, which estimates $\alpha_{\tau,0}(z)$ nonparametrically, points to a rather steep relationship between the house price and its distance to the mine when the house is located in a close vicinity from a mine (smaller values of z) with a diminishing gradient that ultimately plateaus at around a 10-mile mark (just above $z = 50$ thousand feet). Such a shape is remarkably consistent with one's expectation that the property-value effects of environmental disamenities are a *local* phenomenon and that rock mines would not impact values of *distant* properties (with larger values of z). The latter can also be seen from Figure 2, which graphs the estimated gradient of the intercept function along with its 95% confidence bounds. The figure is indicative of a significantly positive effect of z on the log house price within roughly a 10-mile radius of the mine that eventually decreases to a largely statistically insignificant gradient. Actually, as can be seen in Figure 2, for a small share of houses in our sample lying outside the 10-mile radius, the gradient happens to be (statistically) significantly negative implying that, among the farthest houses, those located closer to a rock mine may have higher values, although the estimated effects are timid and not that economically significant.[12] Property-value-boosting employment accessibility effects associated with environmental disamenities may provide a plausible explanation to these findings (see Cohen & Coughlin, 2008).[13] Such affects

[11] Hereafter, in order to facilitate comparability of the results across models, we report the estimates from the logarithmic SQAR model [parametric specification (ii)] in the form of implied semi-elasticities relating the log of house price and the level, not the log, of distance to a nearby rock mine.

[12] With the corresponding median total marginal effect of just –0.08% per 1,000 feet vs. the median total effect of 0.58% per 1,000 feet within the 10-mile radius from the mine (also see, Table 2).

[13] We thank an anonymous referee for bringing this point to our attention.

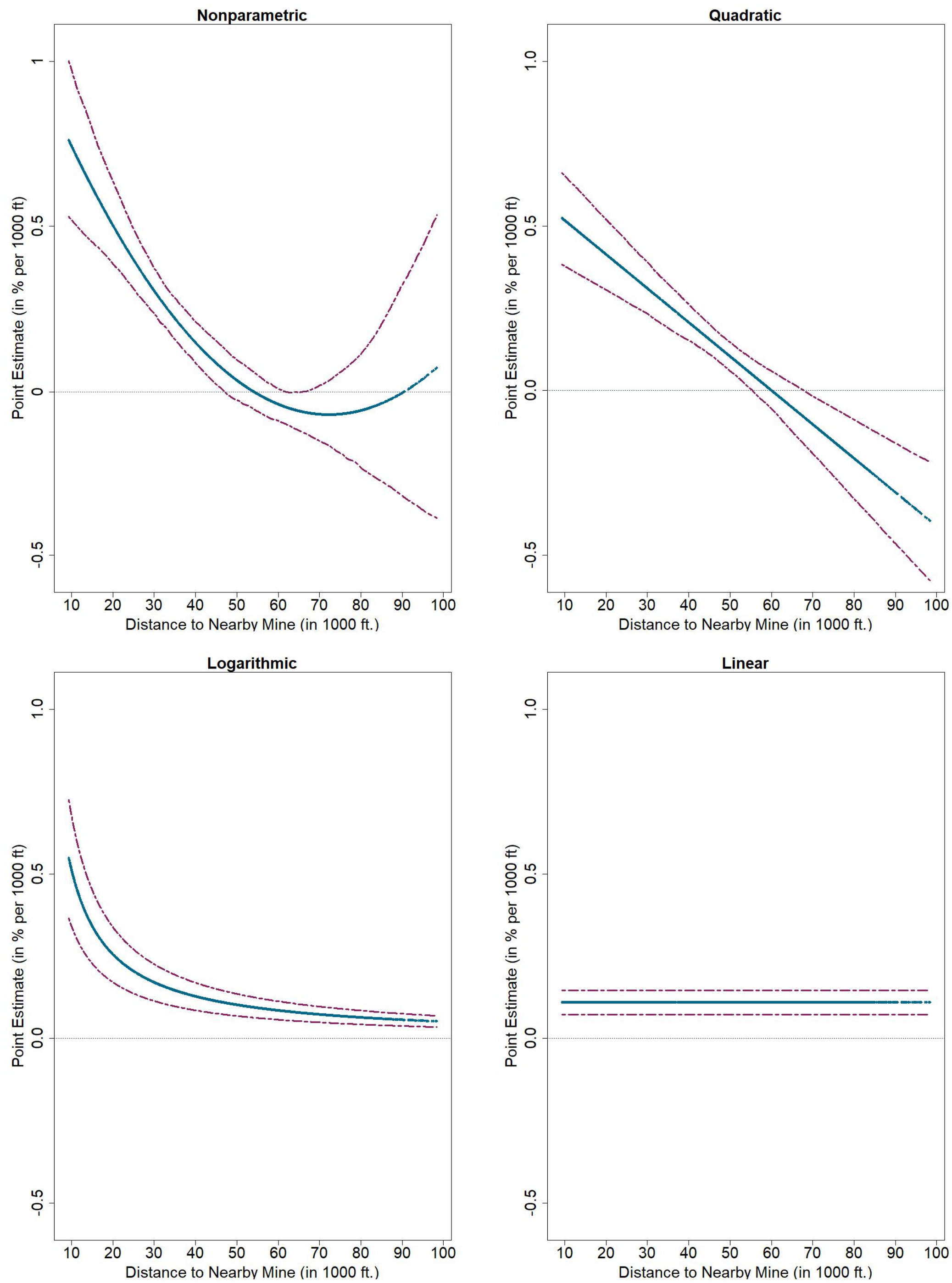

Figure 2. Estimated Gradients of Intercept Functions of the Distance to Rock Mine for the Conditional Median Model (with the 95% bootstrap confidence bounds)

would counteract adverse environmental externality effects of rock mining, implying that the large property-value-suppressing disamenity effects of the rock mining that we find within the 10-mile radius are likely *under*estimated thereby providing further support to their economic significance.

Comparing our model to its parametric alternatives, we expectedly find that parametric models are more susceptible to a functional-form misspecification. While both the quadratic and logarithmic models do successfully find a decreasing gradient of $\alpha_{\tau,0}(z)$ in a close proximity from the mine, neither one is however unable to detect that rock mines appear to become rather irrelevant for the (median) price of houses lying outside their 10-mile radius zone. In fact, a parabolic relationship estimated by the quadratic model rather counter-intuitively suggests a negative (and statistically significant) relationship between the two for large values of z [see Figures 1 and 2]. The logarithmic model does not produce such dubious results but instead implies sample-wise globally significant effects even for unrealistically large distances (including beyond those observable in the data) which is at odds with the local nature of environmental externalities. This illustrates the sensitivity of parametric models (due to their inflexibility) to the inclusion of data on properties that are located farther from the disamenities and thus are hardly, if at all, impacted by negative environmental externalities they generate. To avoid such problems, researchers employing parametric specifications therefore usually have to prespecify a spatial radius of potential impact around the disamenity (e.g., Nelson et al., 1992; Reichert et al., 1992; Hite et al., 2001). However, such an *a priori* choice of the radius is oftentimes *ad hoc* in nature; whereas our model, owing to its nonparametric approach to modeling the distance to disamenity, essentially detects the radius of non-zero impact directly from the data. Lastly, the linear SQAR model fits a linear relationship that is characterized by a rather misleading "average" gradient. The latter can be vividly seen in Figure 2 which shows that, due to its inherent inability to allow for nonlinearities and hence heterogeneity across units, the linear SQAR model tends to grossly under/over-estimate the gradient for nearby/faraway properties.

However, the gradient estimates of $\alpha_{\tau,0}(z)$ plotted in Figure 2 cannot be interpreted as representing marginal partial effects of z on (median) house prices due to the appearance of spatial lag of house prices on the right-hand side of the estimated quantile function. Hence, to obtain partial effects, we consider a reduced form of the fitted outcome variable at the τth quantile: $\widehat{\mathbf{y}}_\tau = [\mathbf{I}_n - \widehat{\rho}_\tau \mathbf{W}_n]^{-1}\left(\mathbf{X}_n\widehat{\boldsymbol{\beta}}_\tau + \widehat{\boldsymbol{\alpha}}_\tau(\mathbf{Z}_n)\right)$, from where we have the following $n \times n$ matrices of marginal effects (MEs):

$$\frac{\partial \widehat{\mathbf{y}}_\tau}{\partial \mathbf{Z}_n'} = [\mathbf{I}_n - \widehat{\rho}_\tau \mathbf{W}_n]^{-1} \times \operatorname{diag}\left\{\frac{\partial \widehat{\alpha}_\tau(z_{1,n})}{\partial z_{1,n}}, \dots, \frac{\partial \widehat{\alpha}_\tau(z_{n,n})}{\partial z_{n,n}}\right\}, \tag{5.1}$$

$$\frac{\partial \widehat{\mathbf{y}}_\tau}{\partial \mathbf{x}_{j,n}'} = [\mathbf{I}_n - \widehat{\rho}_\tau \mathbf{W}_n]^{-1} \times \widehat{\beta}_{\tau,j} \qquad \forall\ j = 1, \dots, d_x, \tag{5.2}$$

where $\mathbf{x}_{j,n} = (x_{j,1}, \dots, x_{j,n})'$ is the jth column of $\mathbf{X}_n$. In the spirit of LeSage & Pace (2009), we refer to the diagonal elements of the gradient matrices of $\widehat{\mathbf{y}}_\tau$ in (5.1)–(5.2) as direct marginal effects (DMEs) and to the off-diagonal elements as indirect marginal effects (IMEs). We analyze marginal effects row-by-row which implies a "*to* a house" interpretation, i.e., how the change in a given covariate across *all* houses affects the price of the ith house. Hence, the summation of elements in the ith row of the gradient matrices in (5.1)–(5.2) provides a measure of the total marginal effect (TME) on the ith house. Also note that, because by design the maximum-eigenvalue-standardized k-nearest-neighbor spatial weights matrix employed in the estimation is in fact row-stochastic, TMEs of covariates that have *constant* gradients (i.e., all "x" variables and, in the case of a linear parametric SQAR model, also variable z) are the same across all observations and are equal to the corresponding gradient times $(1 - \widehat{\rho}_\tau)^{-1}$.

Table 2. Summary of Statistically Significant Point Estimates of ME of the Distance to Rock Mine on Conditional Median of Property Value

	Entire Sample			*Within 10-Mile Radius*		
	TME	DME	IME	TME	DME	IME
			Nonparametric			
5th Perc.	–0.0853	–0.0597	–0.0257	0.1192	0.0831	0.0363
25th Perc.	0.1477	0.1037	0.0433	0.2946	0.2030	0.0885
50th Perc.	0.4629	0.3243	0.1396	0.5810	0.4046	0.1751
75th Perc.	0.8023	0.5581	0.2403	0.8560	0.5957	0.2575
95th Perc.	1.0740	0.7520	0.3227	1.0793	0.7566	0.3245
Mean	0.4836	0.3379	0.1456	0.5768	0.4031	0.1737
			Quadratic			
5th Perc.	–0.3221	–0.2263	–0.0943	0.1271	0.0897	0.0372
25th Perc.	–0.1506	–0.1071	–0.0439	0.2044	0.1449	0.0599
50th Perc.	0.1836	0.1300	0.0535	0.4338	0.3065	0.1272
75th Perc.	0.5108	0.3572	0.1508	0.6130	0.4332	0.1789
95th Perc.	0.7199	0.5063	0.2110	0.7395	0.5226	0.2167
Mean	0.1964	0.1386	0.0577	0.4146	0.2929	0.1217
			Logarithmic			
5th Perc.	0.0938	0.0649	0.0285	0.1460	0.1015	0.0447
25th Perc.	0.1116	0.0773	0.0342	0.1625	0.1128	0.0497
50th Perc.	0.1473	0.1021	0.0450	0.2434	0.1688	0.0744
75th Perc.	0.2532	0.1752	0.0774	0.3975	0.2762	0.1216
95th Perc.	0.5831	0.4038	0.1779	0.7285	0.5042	0.2230
Mean	0.2208	0.1531	0.0677	0.3142	0.2179	0.0964
			Linear			
5th Perc.	0.1646	0.1113	0.0505	0.1646	0.1113	0.0506
25th Perc.	0.1646	0.1124	0.0508	0.1646	0.1113	0.0506
50th Perc.	0.1646	0.1131	0.0514	0.1646	0.1131	0.0515
75th Perc.	0.1646	0.1137	0.0521	0.1646	0.1137	0.0522
95th Perc.	0.1646	0.1140	0.0533	0.1646	0.1140	0.0533
Mean	0.1646	0.1129	0.0516	0.1646	0.1129	0.0517

The reported estimates are in % per 1,000 ft.

The point estimates of total, direct and indirect marginal effects of the (the level of) distance to nearby mine onto the median (log) house price across the four models are summarized in Table 2. Given that insignificant estimates are statistically indistinguishable from zero (implying no effect), here and henceforth, we focus on statistically significant estimates of marginal effects only. For inference within each model, we use the 95% bootstrap percentile confidence bounds.[14] As expected, the results are starkly different across the models, with parametric specifications consistently under-estimating the magnitude of marginal effects of the distance to rock mine on the property value. When considering the entire sample, we find that, in part due to the presence of a large number of houses for which negative marginal effects were estimated, the quadratic model produces estimates of marginal effects on median house values that, on average, are about 59% smaller than those obtained from our semiparametric PLSQAR model. Despite yielding only positive estimates, the logarithmic SQAR model also under-estimates marginal effects, on average, by 54%. The results from a linear model are even more timid (smaller by 66%, on average). Focusing on the more

[14]We use 499 bootstrap replications throughout.

Table 3. Summary of Statistically Significant Semiparametric Estimates of ME of the Distance to Rock Mine on Conditional Quantiles of Property Value within 10-Mile Radius

	TME	DME	IME	TME	DME	IME
	0.25th Q. of Property Value			**0.75th Q. of Property Value**		
25th Perc.	0.3252	0.2182	0.1057	0.3565	0.2676	0.0887
50th Perc.	0.4781	0.3221	0.1571	0.6788	0.5105	0.1688
75th Perc.	0.5645	0.3803	0.1839	0.9979	0.7457	0.2491
Mean	0.4442	0.2993	0.1450	0.6493	0.4875	0.1618
	0.50th Q. of Property Value			**0.95th Q. of Property Value**		
25th Perc.	0.2946	0.2030	0.0885	0.5150	0.3893	0.1256
50th Perc.	0.5810	0.4046	0.1751	0.9952	0.7505	0.2437
75th Perc.	0.8560	0.5957	0.2575	1.3304	1.0048	0.3268
Mean	0.5768	0.4031	0.1737	0.9739	0.7354	0.2385

Reported are the estimates (in % per 1,000 ft) from the PLSQAR model.

economically relevant results confined to a 10-mile radius zone around rock mines, we find that our PLSQAR model suggests the average TME of the distance to the mine on median house prices at around 0.57% per 1,000 feet, 0.40% points of which are the direct effect. The parametric models however yield significantly smaller estimates with the average TMEs of about 0.42% [quadratic], 0.31% [logarithmic] and 0.16% [linear] per 1,000 feet, which are respectively 28%, 46% and 71% smaller than their nonparametric counterpart. The marked difference across our semiparametric model and its three parametric alternatives is apparent not only at the average values of marginal effects but along their entire distributions across houses.

Our comparison of the results from the proposed semiparametric model and those from its parametric counterparts, until now, have largely been casual. However, given that all three parametric specifications are the special cases of our PLSQAR model, we can formally discriminate between the models by means of a specification test described in Section 3. Namely, each parametric median SQAR model can be cast as a restricted model under the null of the first type H_0(i) given in (3.2) to be tested against our unrestricted PLSQAR model. We reject the null in favor of our proposed model in both cases with the bootstrap p-value no larger than 0.032. We also entertain a fourth specification for the parametric SQAR model whereby $\alpha_{\tau,0}(z) \equiv \alpha_{0,\tau}$ for all z, which effectively assumes that z is an irrelevant hedonic attribute that has no effect on the house price. This "constant in z" model serves an auxiliary purpose and is estimated solely in order to facilitate the test of overall relevancy of the house's proximity to a rock mine for its value. In terms of the types of null hypothesis described in Section 3, this restricted model falls under the second type of nulls H_0(ii) given in (3.3), which we test against our PLSQAR model. The corresponding bootstrap p-value is 0.038 suggesting that the proximity to rock mines *does* matter for residential property values.

Given the data lend strong support to our more flexible semiparametric model of house prices, in what follows, we therefore report the results from our PLSQAR model only. Furthermore, in the light of our earlier findings, we focus on the results confined to a local 10-mile radius zone around the mine (2,956 observations) which appear to be the most economically relevant.[15]

Table 3 summarizes statistically significant (house-specific) point estimates of marginal effects of the distance to nearby rock mine on the 0.25th, 0.50th, 0.75th and 0.95th conditional quantiles of the house price from our PLSQAR model. (We caution the reader against confusing quantiles τ

[15] To improve accuracy and to achieve better convergence rates, we still use the full sample during the estimation.

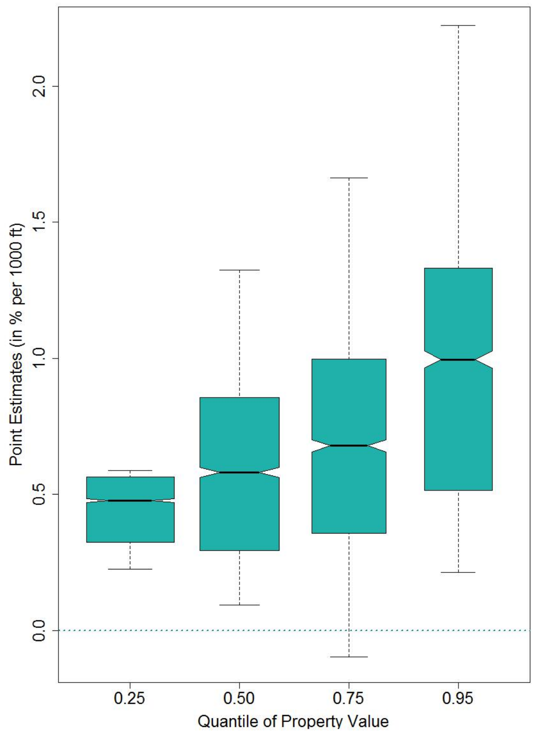

Figure 3. Statistically Significant Semiparametric Estimates of TME of the Distance to Rock Mine on Conditional Quantiles of Property Value within 10-Mile Radius across Quantiles

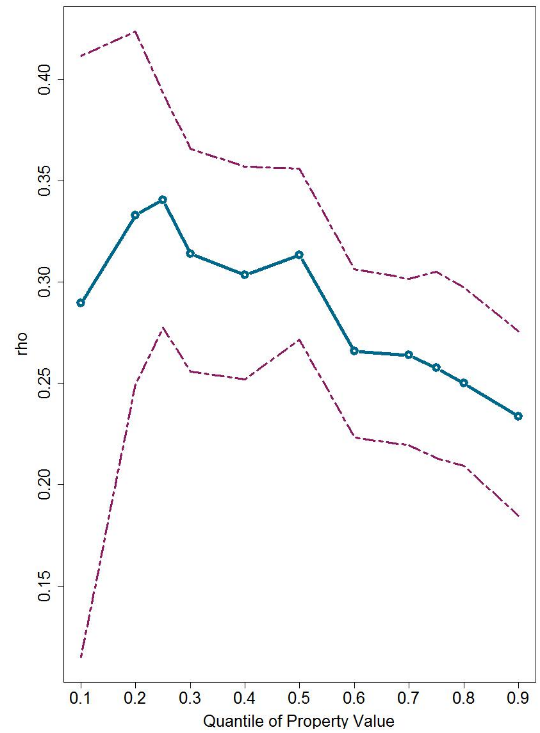

Figure 4. Semiparametric Estimates of the SAR Parameter across Quantiles (with the 95% bootstrap confidence bounds)

of the house price distribution for which model is estimated with quantiles of the fitted distribution of observation-specific marginal effects for *each* τ.) By looking at different quantiles of the house value distribution, we are able to investigate the potentially heterogeneous impact of rock mining on residential property of *different values* thereby looking beyond the results for properties of a "typical" value delivered by standard conditional mean models. Given the tendency of quantile models to be noisier when fitted far in the tails of the distribution, in our analysis we therefore primarily focus on the interquartile range of the conditional house price distribution (setting $\tau = \{0.25, 0.50, 0.75\}$) which should give us sufficient insights into distributional effects, if any, of rock mines on house prices. That said, motivated by the proposition oftentimes claimed in the literature whereby environmental disamenities have significantly larger effects on expensive upscale properties (Reichert et al., 1992; Gayer, 2000), we also estimate our model at the 0.95th quantile to examine if the negative effects of rock mines are especially amplified when located near the most expensive houses. Overall, the results in Table 3 lend strong support to heterogeneous distributional value-suppressing effects of rock mines on the prices of nearby houses, the magnitude of which increase with the value of these houses, as expected. This distributional heterogeneity in the marginal effects can be seen even more vividly in Figure 3 which plots the distribution of the TME estimates across quantiles of the house price distribution. The figure also points to an increase in variability (i.e., a higher degree of heterogeneity across individual houses) of the TME estimates as house prices rise.

As we move from the first to third quartile of the house price distribution, we find that the average estimate of TME of the distance to nearby rock mine on house prices significantly increases from 0.44% to 0.65% per 1,000 feet [see Table 3]. When we focus on the most expensive properties at the 0.95th quantile, the TME goes up even further with the corresponding *median* estimate of about 1% and a half of point estimates being even larger than that; the mean estimate is 0.97% per 1,000 feet. For residential property in the middle of the price distribution ($\tau = 0.50$), our estimates suggest that, between two identical houses, the one located a mile closer to a rock mine is predicted to be priced, on average, at about 3.1% discount.[16] The analogous average discounts for houses in the first and third quartiles of price distribution are around 2.3 and 3.4%, respectively. For upscale property in the 0.95th quantile, it is at an astounding 5.1%. This is rather expected because of income sorting whereby higher income households have higher ability to pay for better environmental quality: in this case, distance from a disamenity. Conversely, households with lower incomes and less expensive homes are perhaps more willing to substitute environmental quality for other, more necessary, house characteristics such as easier access to employment including jobs in the environmental-externality-generating rock mining industry itself.[17] As a back-of-the-envelope welfare calculation using unconditional sample quantiles of house values corresponding to the fitted quantile functions,[18] the above discount estimates imply the average loss in property value associated with the house being located a mile closer to a rock mine ranging from \$3,691 to \$10,970 for houses within the interquartile range of price distribution. For more expensive neighborhoods in the 0.95th quantile, such losses can be, on average, as high as \$28,410. We can further extend the welfare analysis to obtain aggregate property value losses due to the houses' proximity to rock mine by applying the estimated discounts to actual house prices at each observation in order to predict increase in each property's value if it were moved from its actual location to a (counterfactual) 10-mile distance from the mine. Applying this method to properties with statistically significant total marginal effects[19] of the distance lying within a 10-mile radius from the mine, we find a total property value loss of \$68.4 million at the median, which would have a practically substantive impact on public goods expenditures in the county, especially on schools, because of the lost tax revenue amounting to approximately \$1.3 million per annum.

Our estimates of marginal effects also indicate a decreasing (relative) importance of IMEs for residential properties of higher values. While the indirect effects working through neighbors, on average, contribute 37.8% to the TME of z_i on the log house price at the first quartile of the property value distribution, their average contribution falls quite dramatically to 26.6% for the houses at the third quartile. A plausible explanation for this is that less expensive properties may have very different interior quality levels resulting in more unobserved heterogeneity as compared to higher priced houses. Thus, in more expensive neighborhoods, the adverse effects of nearby rock mines are "priced in" directly during the valuation as opposed to via a spillover comparison to neighboring properties. In other words, we find that spatial dependence in house prices decreases as the value of property rises. To see this, consider the estimates of spatial autoregressive parameter which measures spatial dependence in the data. We summarize the estimates of $\rho_{\tau,0}$, along with their confidence bounds, across different τ of the conditional house price distribution in Figure 4. It is evident that the SAR coefficient declines as we move from the left to the right tail of the distribution implying that neighborhood effects are more pronounced in less expensive areas.

[16] 5.28 thousand feet times the mean estimate of 0.58% per 1,000 feet. The average discount estimates for other quantiles of house price are obtained similarly.

[17] Cohen & Coughlin (2008) discuss such positive employment accessibility effects associated with environmental disamenities which may counteract negative externality effects in the context of a noise-generating airport.

[18] And assuming a constant marginal willingness to pay.

[19] Thereby conservatively assuming that the value of houses with insignificant marginal effects of the distance wound not increase.

This result is similar to Liao & Wang's (2012), who estimate a fully parametric hedonic quantile model (however, with no environmental disamenities considered) and also find that the spatial autoregressive parameter declines between the 30th and 70th quantiles. Nonetheless, our estimated spatial effects are statistically significant throughout the entire house price distribution thereby indicating that the failure to account for spatial dependence, as usually done in the literature on housing-market-based valuations of adverse effects of environmental disamenities, would likely yield inconsistent estimates. This substantiates our spatial-econometric approach to hedonic modeling.

6 Conclusion

This paper provides the first estimates of the effects of rock mining—an environmental disamenity—on local residential property values. We estimate the relationship between a house's price and its distance from nearby rock mine in Delaware County, Ohio. We improve upon the conventional approach to valuating adverse effects of environmental disamenities based on hedonic house price functions by developing a novel (semiparametric) partially linear spatial quantile autoregressive model which accommodates unspecified nonlinearities, distributional heterogeneity as well as provides a means to indirectly control for unobservable house and neighborhood characteristics using the spatial dependence in the data. Our model constitutes a practically useful fusion of semi/nonparametric quantile methods with models of spatial dependence. We estimate it via a two-step nonparametric sieve IV quantile estimator. We also propose a model specification test.

We find statistically and economically significant property-value-suppressing effects of being located near an operational rock mine which gradually decline to insignificant near-zero values at a roughly ten-mile distance. Our estimates suggest that, *ceteris paribus*, a house located a mile closer to a rock mine is priced, on average, at about 2.3–5.1% discount, with more expensive properties being subject to larger markdowns. As a back-of-the-envelope welfare calculation, the above discount estimates imply the average loss in property value associated with the house being located a mile closer to a rock mine ranging from $3,691 to $10,970 for houses within the interquartile range of price distribution. For more expensive neighborhoods in the 0.95th quantile, such losses can be, on average, as high as $28,410. Applying the estimated statistically significant discounts to house prices at each observation lying within a 10-mile radius from the mine to predict an increase in each property's value if it were moved from its actual location to a (counterfactual) 10-mile distance from the mine, we find the aggregate property value loss associated with rock mining in the area to be $68.4 million at the median.

Appendix

A Mathematical Proofs

For any $x \neq 0$ and y, we have

$$\zeta_\tau(x-y) - \zeta_\tau(x) = y\varphi_\tau(y) + \int_0^y \left(\mathbb{1}\{x \leq t\} - \mathbb{1}\{x \leq 0\}\right) dt, \tag{A.1}$$

where $\varphi_\tau(u) = \tau - \mathbb{1}\{u < 0\}$. In addition, throughout the Appendix, we define $\mathbf{v}_n(\rho) = [\mathbf{I}_n + (\rho_{\tau,0} - \rho)\,\mathbf{G}_n]\,\mathbf{u}_n$ and the cdf and pdf of its ith element, $v_{i,n}(\rho)$, is denoted by $F_{v_{i,n}(\rho)}(v)$ and $f_{v_{i,n}(\rho)}(v)$, respectively. Below, we first introduce a lemma used to derive our theorems.

Lemma 1 *(i) Under Assumption* ***3****, we have* $\sup_{n,\boldsymbol{l}(i)\in D_n}\sum_{\boldsymbol{l}(j)\in D_n,\varrho(i,j)>s}|g_{ij,n}|\leq Ms^{-c_2d}$*; (ii) under Assumptions* ***2–3****,* $\{v_{i,n}(\rho),\boldsymbol{l}(i)\in D_n\}$ *is uniformly* L_2*-NED on* $\{\varepsilon_{in},\boldsymbol{l}(i)\in D_n\}$ *with the NED coefficients of* $\psi(s)=O(s^{-\varsigma})$*; (iii)* $f'_{v_{i,n}(\rho)}(v)$ *is continuous and uniformly bounded with respect to* $v\in R^n$ *across index* i *and*

$$\frac{1}{n}\sum_{i=1}^{n}\left\{f_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)-\mathbb{E}\left[f_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)\right]\right\}\mathcal{X}_{i,n}\mathcal{X}'_{i,n}\ \xrightarrow{p}\ \mathbf{0}_{1+d_x+d_m+L_n},\tag{A.2}$$

where $\eta_{i,n}(\rho)=(\rho-\rho_{\tau,0})\sum_{j=1}^{n}g_{ij,n}\left[\mathbf{x}'_{j,n}\beta_{\tau,0}+\alpha_{\tau,0}(\mathbf{z}_{j,n})\right]+\mathbf{x}'_{i,n}\left[\boldsymbol{\beta}_{\tau,0}(\rho)-\boldsymbol{\beta}_{\tau,0}\right]+\alpha^*_{\tau,0}(\mathbf{z}_{i,n},\rho)-\alpha_{\tau,0}(\mathbf{z}_{i,n})+\mathbf{m}'_{i,n}\gamma_{\tau,0}(\rho)$ *and* $\alpha^*_{\tau,0}(\mathbf{z}_{i,n},\rho)=\boldsymbol{\phi}_{L_n}(\mathbf{z}_{i,n})'\mathcal{A}_{\tau,0}(\rho)$*.*

Proof. (i) Under Assumption **3**, we have $\mathbf{G}_n\equiv\mathbf{W}_n\mathbf{S}_n^{-1}=\mathbf{W}_n\sum_{k=0}^{\infty}(\rho_{\tau,0}\mathbf{W}_n)^k=\mathbf{W}_n+\rho_{\tau,0}\mathbf{W}_n^2+\rho_{\tau,0}^2\mathbf{W}_n^3+\ldots$, and hence we have

$$\begin{aligned}g_{ij,n}\ &=\ w_{ij,n}+\rho_{\tau,0}\sum_{l\neq i}w_{il,n}w_{lj,n}+\rho_{\tau,0}^2\sum_{l_2\neq i}w_{il_2,n}\left(\sum_{l_1\neq l_2}w_{l_2l_1,n}w_{l_1j,n}\right)+\rho_{\tau,0}^3\sum_{l_3\neq i}w_{il_3,n}\sum_{l_2\neq l_3}w_{l_3l_2,n}\\&\quad\times\left(\sum_{l_1\neq l_2}w_{l_2l_1,n}w_{l_1j,n}\right)+\ldots+\rho_{\tau,0}^k\sum_{l_k\neq i}\sum_{l_{k-1}\neq l_k}\cdots\sum_{l_1\neq l_2}w_{il_k,n}w_{l_kl_{k-1},n}\ldots w_{l_2l_1,n}w_{l_1j,n}+\ldots.\end{aligned}$$

For all j such that $\boldsymbol{l}(j)\in D_n$ and $\varrho(i,j)>s$, we have

$$\begin{aligned}\sum_{\boldsymbol{l}(j)\in D_n,\varrho(i,j)>s}|g_{ij,n}|\ &\leq\ c_1\sum_{k=1}^{\infty}\rho_{\tau,0}^{k-1}\left(\frac{s}{k}\right)^{-c_2d}\leq\frac{c_1s^{c_2d}}{\rho_{\tau,0}}\int_1^{\infty}\rho_{\tau,0}^{x}x^{[c_2d]+1}dx\\&=\ -\frac{c_1s^{-c_2d}}{(\ln\rho_{\tau,0})^{[c_2d]+2}}\sum_{k=0}^{[c_2d]+1}\frac{([c_2d]+1)!}{([c_2d]+1-k)!}(-\ln\rho_{\tau,0})^{[c_2d]+1-k},\end{aligned}$$

where $[a]$ is the largest integer smaller than $a>0$. This completes the proof of (i).

(ii) By definition, $v_{i,n}(\rho)=u_{i,n}+(\rho_{\tau,0}-\rho)\sum_{j=1}^{n}g_{ij,n}u_{j,n}$. Applying Minkowski's and conditional Jensen's inequalities yields

$$\begin{aligned}\|v_{i,n}(\rho)-\mathbb{E}[v_{i,n}(\rho)|\mathcal{F}_{i,n}(s)]\|_2\ &\leq\ \|u_{i,n}-\mathbb{E}[u_{i,n}|\mathcal{F}_{i,n}(s)]\|_2+|\rho_{\tau,0}-\rho|\sum_{j=1}^{n}|g_{ij,n}|\,\|u_{j,n}-\mathbb{E}[u_{j,n}|\mathcal{F}_{i,n}(s)]\|_2\\&\leq\ M\psi(s)+2|\rho_{\tau,0}-\rho|\sum_{\{j:\varrho(i,j)>s\}}|g_{ij,n}|\,\|u_{j,n}\|_2.\end{aligned}$$

This completes the proof of (ii).

(iii) Let $f_{\mathbf{v}_n(\rho)}(v)$ be the joint pdf of $\mathbf{v}_n(\rho)$. Then, we have $f_{\mathbf{v}_n(\rho)}(v)=|\mathbf{I}_n+(\rho_{\tau,0}-\rho)\mathbf{G}_n|^{-1}\times f_{\mathbf{u}_n}\left([\mathbf{I}_n+(\rho_{\tau,0}-\rho)\mathbf{G}_n]^{-1}\mathbf{v}_n(\rho)\right)$. Then, under Assumption **3(i)–(ii)**, we see that $f_{\mathbf{v}_n(\rho)}(v)$ is continuously differentiable with respect to each element of $v\in R^n$. Applying the Leibniz rule under Assumption **5(i)**, we have that $f'_{v_{i,n}(\rho)}(v)$ is continuous and uniformly bounded with respect to $v\in R^n$ across index i. Lastly, applying Theorem 1 in Jenish & Prucha (2012) yields (A.2). ■

Proof of Theorem 1. Denote $\widehat{\boldsymbol{\vartheta}}_\tau(\rho)=\sqrt{n}\left[\widehat{\boldsymbol{\theta}}_\tau(\rho)-\boldsymbol{\theta}_{\tau,0}(\rho)\right]$, $Y^*_{i,n}(\rho)=y_{i,n}-\rho\sum_{j\neq i}w_{ij,n}y_{j,n}-\mathcal{X}'_{i,n}\boldsymbol{\theta}_{\tau,0}(\rho)=v_{i,n}(\rho)-\eta_{i,n}$ and $Y_{i,n}(\rho)=Y^*_{i,n}(\rho)-n^{-1/2}\mathcal{X}'_{i,n}\boldsymbol{\vartheta}_\tau(\rho)$. Then, for any given $\rho\in\Lambda_\rho$, solving the optimization problem (2.9) in Step 1 is equivalent to finding $\widehat{\boldsymbol{\vartheta}}_\tau(\rho)$ that minimizes

$$Q_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)=\frac{1}{n}\sum_{i=1}^{n}\left(\zeta_\tau\left(Y_{i,n}(\rho)\right)-\zeta_\tau\left(Y_{i,n}^*(\rho)\right)\right), \tag{A.3}$$

which is convex in $\boldsymbol{\vartheta}_\tau(\rho)$. We can show that, under Assumptions **2** and **5**,

$$\Pr\left[\sum_{i=1}^{n}\mathbb{1}\left\{Y_{i,n}^*(\rho)=0\right\}=O(1)\right]=1 \quad \text{almost surely over all } \rho\in\Lambda_\rho. \tag{A.4}$$

Following the proof method used in Fan et al. (1994), we have

$$Q_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)=\mathbb{E}\left[Q_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)\right]+\frac{\boldsymbol{\vartheta}_\tau(\rho)'}{n^{3/2}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\left(\varphi_\tau\left(Y_{i,n}^*(\rho)\right)-\mathbb{E}\left[\varphi_\tau\left(Y_{i,n}^*(\rho)\right)\right]\right)+R_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right),$$

where $R_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)=n^{-1}\sum_{i=1}^{n}\left(Q_{i,n}(\rho)-\mathbb{E}\left[Q_{i,n}(\rho)\right]\right)$ is a dominated term in the large sample as shown below, and

$$\begin{aligned}
Q_{i,n}(\rho) &= \zeta_\tau\left(Y_{i,n}(\rho)\right)-\zeta_\tau\left(Y_{i,n}^*(\rho)\right)-n^{-1/2}\boldsymbol{\vartheta}_\tau(\rho)'\mathcal{X}_i\varphi_\tau\left(Y_{i,n}^*(\rho)\right)\\
&= \int_0^{t_{i,n}}\left[\mathbb{1}\left\{v_{i,n}(\rho)\le\eta_{i,n}(\rho)+t\right\}-\mathbb{1}\left\{v_{i,n}(\rho)\le\eta_{i,n}(\rho)\right\}\right]dt.
\end{aligned}$$

Denoting $t_{i,n}=n^{-1/2}\mathcal{X}_{i,n}'\boldsymbol{\vartheta}_\tau(\rho)$ and applying (A.1) and (A.4), we obtain

$$\begin{aligned}
\mathbb{E}\left[Q_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\zeta_\tau\left(Y_{i,n}^*(\rho)-n^{-1/2}\mathcal{X}_{i,n}'\boldsymbol{\vartheta}_\tau(\rho)\right)-\zeta_\tau\left(Y_{i,n}^*(\rho)\right)\right]\\
&\approx \frac{\boldsymbol{\vartheta}_\tau(\rho)'}{n^{3/2}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\mathbb{E}\left[\varphi_\tau\left(Y_{i,n}^*(\rho)\right)\right]+\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_{i,n}}\mathbb{E}\left[\mathbb{1}\left\{Y_{i,n}^*(\rho)\le t\right\}-\mathbb{1}\left\{Y_{i,n}^*(\rho)\le 0\right\}\right]dt\\
&= \frac{\boldsymbol{\vartheta}_\tau(\rho)'}{n^{3/2}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\mathbb{E}\left[\varphi_\tau\left(Y_{i,n}^*(\rho)\right)\right]+\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_{i,n}}\left[F_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)+t\right)-F_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)\right]dt\\
&= \frac{\boldsymbol{\vartheta}_\tau(\rho)'}{n^{3/2}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\mathbb{E}\left[\varphi_\tau\left(Y_{i,n}^*(\rho)\right)\right]+\\
&\quad \frac{\boldsymbol{\vartheta}_\tau(\rho)'}{2n^2}\sum_{i=1}^{n}f_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)\mathcal{X}_{i,n}\mathcal{X}_{i,n}'\ \boldsymbol{\vartheta}_\tau(\rho)+O_p\left(\left(\frac{L_n}{\sqrt{n}}\right)^{3/2}\right),
\end{aligned}$$

where $F_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)+t\right)-F_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)=f_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)t+f_{v_{i,n}(\rho)}'\left(\bar{\eta}_{i,n}(\rho)\right)t^2/2$ with $\bar{\eta}_{i,n}(\rho)$ lying between $\eta_{i,n}(\rho)$ and $\eta_{i,n}(\rho)+t$, and

$$\left|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_{i,n}}f_{v_{i,n}(\rho)}'\left(\bar{\eta}_{i,n}(\rho)\right)t^2dt\right|\le\frac{M}{3n^{5/2}}\sum_{i=1}^{n}\left|\mathcal{X}_{i,n}'\boldsymbol{\vartheta}_\tau(\rho)\right|^3=O_p\left(n^{-3/2}L_n^3\right)=o_p(1)$$

by Lemma **1(iii)** under Assumptions **5–6**.

Next, we consider $R_n\left(\boldsymbol{\vartheta}_\tau(\rho)\right)$. Letting $\widetilde{v}_{i,n}(\rho)=\mathbb{E}\left[v_{i,n}(\rho)\,|\mathcal{F}_{i,n}(s)\right]$, $\epsilon_{i,n}(\rho)=v_{i,n}(\rho)-\widetilde{v}_{i,n}(\rho)$ and $\widetilde{Q}_{i,n}(\rho)=\int_0^{t_{i,n}}\left[\mathbb{1}\left\{\widetilde{v}_{i,n}(\rho)\le\eta_{i,n}(\rho)+t\right\}-\mathbb{1}\left\{\widetilde{v}_{i,n}(\rho)\le\eta_{i,n}(\rho)\right\}\right]dt$, we have $Q_{i,n}(\rho)-\widetilde{Q}_{i,n}(\rho)=\int_0^{t_{i,n}}\left[\mathbb{1}\left\{v_{i,n}(\rho)\in A_{i,n}\right\}-\mathbb{1}\left\{v_{i,n}(\rho)\in B_{i,n}\right\}\right]dt$, where $A_{i,n}$ is an interval lying between

$\eta_{i,n}(\rho)$ and $\eta_{i,n}(\rho)+\epsilon_{i,n}(\rho)$ and $B_{i,n}$ is an interval lying between $\eta_{i,n}(\rho)+t$ and $\eta_{i,n}(\rho)+\epsilon_{i,n}(\rho)+t$. We then can show that

$$\begin{aligned} \left\| Q_{i,n}(\rho) - \mathbb{E}\left[Q_{i,n}(\rho)\,|\mathcal{F}_{i,n}(s)\right]\right\|_2^2 &\leq \left\| Q_{i,n}(\rho) - \widetilde{Q}_{i,n}(\rho)\right\|_2^2 + \left\|\mathbb{E}\left[Q_{i,n}(\rho) - \widetilde{Q}_{i,n}(\rho)\,|\mathcal{F}_{i,n}(s)\right]\right\|_2^2 \\ &\leq M\left\|\epsilon_{i,n}(\rho)\right\|_2^2, \end{aligned}$$

because the conditional pdf of $v_{i,n}(\rho)$ given $\epsilon_{i,n}(\rho)$ (and $\mathcal{F}_{i,n}(s)$) is uniformly bounded over index i. By Definition 1 in Jenish & Prucha (2012), we therefore obtain that $\{Q_{i,n}(\rho), \boldsymbol{l}(i)\in D_n\}$ is uniformly L_2-NED on $\{\varepsilon_{i,n}, \boldsymbol{l}(i)\in D_n\}$ with the same NED mixing coefficients as those for $\{v_{i,n}(\rho), \boldsymbol{l}(i)\in D_n\}$. It is readily seen that $\sum_{s=1}^{\infty} s^{d-1}\psi(s) \leq M\sum_{s=1}^{\infty} s^{d-\varsigma-1} < M$ because $\varsigma > d$, and $\mathbb{E}\left[|Q_{i,n}(\rho)|^{2+\delta}\right] \leq \mathbb{E}\left[|t_{i,n}|^{2+\delta}\right] \leq Mn^{-(2+\delta)/2}L_n^{2+\delta} \to 0$ for any $\delta > 0$ as $n\to\infty$ under Assumption **6**. By Lemma A.3(a) in Jenish & Prucha (2012), we obtain $\mathbb{V}\mathrm{ar}\left[R_n(\boldsymbol{\vartheta}_\tau(\rho))\right] \leq ML_n^3/n^{3/2}$ under Assumption **2(iii)**. Hence, we obtain $R_n(\boldsymbol{\vartheta}_\tau(\rho)) = O_p\left((L_n/\sqrt{n})^{3/2}\right)$.

Combining the above results gives

$$Q_n(\boldsymbol{\vartheta}_\tau(\rho)) = \frac{\boldsymbol{\vartheta}_\tau(\rho)'}{n^{3/2}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_\tau\left(Y_{i,n}^*(\rho)\right) + \frac{1}{2n}\boldsymbol{\vartheta}_\tau(\rho)'\ \boldsymbol{\Sigma}_\tau(\rho)\ \ \boldsymbol{\vartheta}_\tau(\rho) + o_p(1) \tag{A.5}$$

under Assumption **6**, and this result holds uniformly over $\rho\in\Lambda_\rho$ by the convexity lemma of Pollard (1991), where

$$\boldsymbol{\Sigma}_\tau(\rho) = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} f_{v_{i,n}(\rho)}\left(\eta_{i,n}(\rho)\right)\mathcal{X}_{i,n}\mathcal{X}_{i,n}' \tag{A.6}$$

by Lemma **1(iii)**. Then, minimizing (A.5) with respect to $\boldsymbol{\vartheta}_\tau(\rho)$ yields

$$\widehat{\boldsymbol{\vartheta}}_\tau(\rho) = -\frac{\boldsymbol{\Sigma}_\tau^{-1}(\rho)}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_\tau\left(Y_{i,n}^*(\rho)\right) + o_p(1), \tag{A.7}$$

which holds uniformly over $\rho\in\Lambda_\rho$. So, we obtain

$$\sqrt{n}\left[\widehat{\boldsymbol{\theta}}_\tau(\rho) - \boldsymbol{\theta}_{\tau,0}(\rho)\right] = -\frac{\boldsymbol{\Sigma}_\tau^{-1}(\rho)}{\sqrt{n}}\sum_{i=1}^{n}\left[\tau - \mathbb{1}\left\{v_{i,n}(\rho)\leq\eta_{i,n}(\rho)\right\}\right]\mathcal{X}_{i,n} + o_p(1). \tag{A.8}$$

Applying Lemma **1(ii)** and the CLT of Jenish & Prucha (2012, Theorem 2), we obtain that $n^{-1/2}\sum_{i=1}^{n}\left(\mathbb{1}\{v_{i,n}(\rho)\leq\eta_{i,n}(\rho)\} - \mathbb{E}\left[\mathbb{1}\{v_{i,n}(\rho)\leq\eta_{i,n}(\rho)\}\right]\right)\mathcal{X}_{i,n} = O_p(1)$ element by element, where $n^{-1}\sum_{i=1}^{n}\left\{\tau - \mathbb{E}\left[\mathbb{1}\{v_{i,n}(\rho)\leq\eta_{i,n}(\rho)\}\right]\right\}\mathcal{X}_{i,n} = 0$ since this term is the first-order condition of $\max_{\boldsymbol{\vartheta}_\tau(\rho)}\mathbb{E}\left[Q_n(\boldsymbol{\vartheta}_\tau(\rho))\right]$. Hence, under Assumption **5(iii)**, we obtain $\left\|\widehat{\boldsymbol{\theta}}_\tau(\rho) - \boldsymbol{\theta}_{\tau,0}(\rho)\right\| = O_p\left(\sqrt{L_n/n}\right)$ uniformly over ρ. This completes the proof of this theorem. ■

Proof of Theorem 2. In Step 2, we calculate $\widehat{\rho}_\tau = \arg\min_\rho\ \widehat{\boldsymbol{\gamma}}_\tau(\rho)'\mathbf{V}_n\widehat{\boldsymbol{\gamma}}_\tau(\rho)$, where $\widehat{\boldsymbol{\gamma}}_\tau(\rho) = \boldsymbol{\gamma}_{\tau,0}(\rho)+o_p(1)$ uniformly over ρ by Theorem **1**. Since $\widehat{\boldsymbol{\gamma}}_\tau(\rho)$ is continuous in ρ and $\boldsymbol{\gamma}_{\tau,0}(\rho)'\mathbf{V}_n\ \boldsymbol{\gamma}_{\tau,0}(\rho)$ has a minimum value at $\rho_{\tau,0}$, we obtain $\widehat{\rho}_\tau \xrightarrow{p} \rho_{\tau,0}$ by Theorem 2.1 in Newey & McFadden (1994). Since $\boldsymbol{\theta}_{\tau,0}(\rho)$ is continuous in ρ, we have $\left\|\widehat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_{\tau,0}\right\| = o_p(1)$.

When $\rho = \rho_{\tau,0}$, we have $v_{i,n}(\rho_{\tau,0}) = u_{i,n}$, $\eta_{i,n}(\rho_{\tau,0}) = \alpha_{\tau,0}^*(\mathbf{z}_{i,n}) - \alpha_{\tau,0}(\mathbf{z}_{i,n})$, and

$$\boldsymbol{\Sigma}_\tau = \boldsymbol{\Sigma}_\tau(\rho_{\tau,0}) = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} f_{u_{i,n}}(0)\,\mathcal{X}_i\mathcal{X}_i' \quad \text{by (2.15) and (A.6).}$$

Let ρ_n be a constant satisfying $\rho_n = \rho_{\tau,0} + o(1)$, and denote $\widehat{Y}_{i,n}(\rho_n) = y_{i,n} - \rho_n \sum_{j \neq i} w_{ij,n} y_{j,n} - \mathcal{X}'_{i,n} \widehat{\boldsymbol{\theta}}_\tau(\rho_n) = Y^*_{i,n}(\rho_n) + \mathcal{X}'_{i,n}\left(\boldsymbol{\theta}_{\tau,0}(\rho_n) - \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)$. By Lemma A.2 in Ruppert & Carroll (1980) , we have $o_p(1) = n^{-1/2}\sum_{i=1}^n \varphi_\tau\left(\widehat{Y}_{i,n}(\rho_n)\right)\mathcal{X}_{i,n}$. Let $\chi_{i,n}(\rho, \boldsymbol{\theta}) = \varphi_\tau\left(y_{i,n} - \rho\sum_{j\neq i} w_{ij,n} y_{j,n} - \mathcal{X}'_{i,n}\boldsymbol{\theta}(\rho)\right)\mathcal{X}_{i,n}$ and $\mathbb{E}\left[\chi_{i,n}\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)\right] = \mathbb{E}\left[\chi_{i,n}(\rho, \boldsymbol{\theta}(\rho))\right]_{(\rho,\boldsymbol{\theta}(\rho)) = \left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)}$ and decompose

$$\begin{aligned} \frac{1}{\sqrt{n}}\sum_{i=1}^n \varphi_\tau\left(\widehat{Y}_{i,n}(\rho_n)\right)\mathcal{X}_{i,n} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n\left[\chi_{i,n}\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right) - \mathbb{E}\left[\chi_{i,n}\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)\right]\right] \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{E}\left[\chi_{i,n}\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)\right]. \end{aligned}$$

First, since $\mathbb{E}\left[\chi_{i,n}(\rho, \boldsymbol{\theta}_\tau(\rho))\right] = \mathbb{E}\left[\tau - \mathbb{1}\{v_{i,n}(\rho) \leq \eta_{i,n}(\rho)\}\right]\mathcal{X}_{i,n}$, we have

$$\begin{aligned} \mathbb{E}\left[\tau - \mathbb{1}\{v_{i,n}(\rho) \leq \eta_{i,n}(\rho)\}\right] &= \mathbb{E}\left[F_{u_{i,n}}(0|\bar{u}_{i,n}) - F_{u_{i,n}}\left(\frac{(\rho_{\tau,0} - \rho)\bar{u}_{i,n} + \eta_{i,n}(\rho)}{1 + (\rho_{\tau,0} - \rho) g_{ii,n}} \,\middle|\, \bar{u}_{i,n}\right)\right] \\ &= -\mathbb{E}\left[\bar{u}_{i,n} f_{u_{i,n}}(\bar{c}_{i,n}|\bar{u}_{i,n})\right]\frac{\rho_{\tau,0} - \rho}{1 + (\rho_\tau - \rho) g_{ii,n}} \\ &\quad - \mathbb{E}\left[f_{u_{i,n}}(\bar{c}_{i,n}|\bar{u}_{i,n})\right]\frac{\eta_{i,n}(\rho)}{1 + (\rho_{\tau,0} - \rho) g_{ii,n}} \end{aligned}$$

if $1 + (\rho_{\tau,0} - \rho) g_{ii,n} > 0$, where $\bar{c}_{i,n}$ lies between 0 and $\left[(\rho_{\tau,0} - \rho)\bar{u}_{i,n} + \eta_{i,n}(\rho)\right]\left[1 + (\rho_{\tau,0} - \rho) g_{ii,n}\right]^{-1}$. Therefore, if $\lim_{n\to\infty}\inf_{1\leq i\leq n}\left[1 + (\rho_{\tau,0} - \rho_n) g_{ii,n}\right] = c_g > 0$, we obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{E}\left[\chi_{i,n}\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right)\right] \approx -\mathcal{A}_1\sqrt{n}(\rho_{\tau,0} - \rho_n) - \mathcal{A}_2\sqrt{n}\left(\boldsymbol{\theta}_{\tau,0}(\rho_n) - \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right),$$

where

$$\begin{aligned} \mathcal{A}_1 &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\left[1 + (\rho_{\tau,0} - \rho_n) g_{ii,n}\right]^{-1}\mathbb{E}\left[f_{u_{i,n}}(0|\bar{u}_{i,n})\left(\bar{u}_{i,n} + \sum_{j=1}^n g_{ij,n}\left[\mathbf{x}'_{j,n}\boldsymbol{\beta}_{\tau,0} + \alpha_{\tau,0}(\mathbf{z}_{j,n})\right]\right)\right]\mathcal{X}_{i,n}, \\ \mathcal{A}_2 &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\left[1 + (\rho_{\tau,0} - \rho_n) g_{ii,n}\right]^{-1}\mathbb{E}\left[f_{u_{i,n}}(0|\bar{u}_{i,n})\right]\mathcal{X}_{i,n}\mathcal{X}'_{i,n}. \end{aligned} \tag{A.9}$$

Second, Jenish (2016) has proven the stochastic equicontinuity result of an empirical process for the smooth function of a NED spatial process and finite parameters. Applying Theorem 5 in Jenish (2016), we obtain that the equicontinuity result also holds for $\Delta_n(\rho, \boldsymbol{\theta}(\rho))$ here, i.e.,

$$\left\|\Delta_n\left(\rho_n, \widehat{\boldsymbol{\theta}}_\tau(\rho_n)\right) - \Delta_n\left(\rho_{\tau,0}, \boldsymbol{\theta}_{\tau,0}(\rho_{\tau,0})\right)\right\| = o_p(1), \tag{A.10}$$

where

$$\mathcal{A}_{n,0} \equiv \Delta_n\left(\rho_{\tau,0}, \boldsymbol{\theta}_{\tau,0}(\rho_{\tau,0})\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\varphi_\tau\left(Y^*_{i,n}(\rho_{\tau,0})\right) - \mathbb{E}\left[\varphi_\tau\left(Y^*_{i,n}(\rho_{\tau,0})\right)\right]\right)\mathcal{X}_{i,n}. \tag{A.11}$$

Third, combining the above results yields

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_\tau(\rho_n) - \boldsymbol{\theta}_{\tau,0}(\rho_n)\right) = -\mathcal{A}_2^{-1}\mathcal{A}_{n,0} + \mathcal{A}_2^{-1}\mathcal{A}_1\sqrt{n}(\rho_{\tau,0} - \rho_n).$$

Partition below matrix/vector conformably with 1, 2 and 3 corresponding to $\mathbf{x}_{i,n}$, $\mathbf{m}_{i,n}$ and $\boldsymbol{\phi}_{L_n}(\mathbf{z}_{i,n})$, respectively:

$$\mathcal{A}_1 = \begin{bmatrix} \mathcal{A}_{1,1} \\ \mathcal{A}_{1,2} \\ \mathcal{A}_{1,3} \end{bmatrix}, \ \mathcal{A}_2^{-1} = \begin{bmatrix} \mathcal{A}_2^{11} & \mathcal{A}_2^{12} & \mathcal{A}_2^{13} \\ \mathcal{A}_2^{21} & \mathcal{A}_2^{22} & \mathcal{A}_2^{23} \\ \mathcal{A}_2^{31} & \mathcal{A}_2^{32} & \mathcal{A}_2^{33} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_2^{1} \\ \mathcal{A}_2^{2} \\ \mathcal{A}_2^{3} \end{bmatrix} \text{ and } \mathcal{A}_{n,0} = \begin{bmatrix} \mathcal{A}_{n,0,1} \\ \mathcal{A}_{n,0,2} \\ \mathcal{A}_{n,0,3} \end{bmatrix}.$$

Then, we have

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\beta}}_\tau(\rho_n) - \boldsymbol{\beta}_{\tau,0}(\rho_n) \\ \widehat{\boldsymbol{\gamma}}_\tau(\rho_n) - \boldsymbol{\gamma}_{\tau,0}(\rho_n) \end{pmatrix} = -\begin{bmatrix} \mathcal{A}_2^1 \\ \mathcal{A}_2^2 \end{bmatrix}\mathcal{A}_{n,0} + \begin{bmatrix} \mathcal{A}_2^1 \\ \mathcal{A}_2^2 \end{bmatrix}\mathcal{A}_1\sqrt{n}\left(\rho_{\tau,0}-\rho_n\right) . \tag{A.12}$$

In addition, from Step 2, we have $\widehat{\rho}_\tau = \arg\min_{\rho_n} \widehat{\boldsymbol{\gamma}}_\tau'(\rho_n)\mathbf{V}_n\widehat{\boldsymbol{\gamma}}_\tau(\rho_n)$. Applying the CLT of Jenish & Prucha (2012) gives

$$\mathbf{e}_j'\mathcal{A}_{n,0} \xrightarrow{d} \mathbb{N}\left(\mathbf{0}, \mathbf{e}_j'\boldsymbol{\Omega}_\tau\mathbf{e}_j\right)$$

with

$$\boldsymbol{\Omega}_\tau = \lim_{n\to\infty} n^{-1}\sum_{i=1}^n\sum_{j=1}^n \mathcal{X}_{i,n}\mathbb{C}\text{ov}\left[\varphi_\tau(u_{i,n}), \varphi_\tau(u_{j,n})\right]\mathcal{X}_{jn}' = O(1)$$

by Lemma A.3 in Jenish & Prucha (2012), where $\mathbf{e}_j$ is any one of the column vectors of $\mathbf{I}_{d_x+d_m+L_n}$. It follows that $\sqrt{n}\left\|\widehat{\boldsymbol{\gamma}}_\tau(\rho_n) - \boldsymbol{\gamma}_{\tau,0}(\rho_n)\right\| = O_p(1)$, which implies that $\sqrt{n}(\rho_{\tau,0}-\rho_n) = O(1)$. Consequently, we obtain

$$\sqrt{n}\left(\widehat{\rho}_\tau - \rho_{\tau,0}\right) = \left[\mathcal{A}_1'\left(\mathcal{A}_2^2\right)' \ \mathbf{V}_n\mathcal{A}_2^2\mathcal{A}_1\right]^{-1}\mathcal{A}_1'\left(\mathcal{A}_2^2\right)'\mathbf{V}_n\mathcal{A}_2^2\mathcal{A}_{n,0} \equiv \mathcal{M}_\rho\mathcal{A}_{n,0}.$$

Therefore we have

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\beta}}_\tau(\rho_n) - \boldsymbol{\beta}_{\tau,0}(\rho_n) \\ \widehat{\boldsymbol{\gamma}}_\tau(\rho_n) - \boldsymbol{\gamma}_{\tau,0}(\rho_n) \end{pmatrix} = \begin{bmatrix} \mathcal{A}_2^1 \\ \mathcal{A}_2^2 \end{bmatrix}\left(\mathcal{A}_1\mathcal{M}_\rho - \mathbf{I}_{d_x+d_m+L_n}\right)\mathcal{A}_{n,0},$$

so that we obtain

$$\sqrt{n}\boldsymbol{\Sigma}_n^{-1/2}\begin{bmatrix} \widehat{\rho}_\tau - \rho_{\tau,0} & \widehat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_{\tau,0}' & \widehat{\boldsymbol{\gamma}}_\tau' \end{bmatrix}' \xrightarrow{d} \mathbb{N}(\mathbf{0}, \mathbf{I}),$$

where $\boldsymbol{\Sigma}_n = \mathcal{P}\boldsymbol{\Omega}_\tau\mathcal{P}'$ with $\mathcal{P} = \left[\mathcal{M}_\rho', \ \boldsymbol{\Psi}'\left[\left(\mathcal{A}_2^1\right)', \left(\mathcal{A}_2^2\right)'\right]\right]'$ and $\boldsymbol{\Psi} = \mathcal{A}_1\mathcal{M}_\rho - \mathbf{I}_{d_x+d_m+L_n}$.

Lastly, we have

$$\sqrt{n}\left[\widehat{\alpha}_\tau(\mathbf{z}) - \alpha_{\tau,0}(\mathbf{z})\right] = \sqrt{n}\boldsymbol{\phi}_{L_n}(\mathbf{z})'\left(\widehat{\mathcal{A}}_\tau - \mathcal{A}_{\tau,0}\right) = \boldsymbol{\phi}_{L_n}(\mathbf{z})'\mathcal{A}_2^3\left[\mathcal{A}_1\mathcal{M}_\rho - \mathbf{I}_{d_x+d_m+L_n}\right]\mathcal{A}_{n,0},$$

and hence obtain

$$\sqrt{n/\omega_{n,\tau}}\left[\widehat{\alpha}_\tau(\mathbf{z}) - \alpha_{\tau,0}(\mathbf{z})\right] \xrightarrow{d} \mathbb{N}(0,1),$$

where $\omega_{n,\tau} = \boldsymbol{\phi}_{L_n}(\mathbf{z})'\mathcal{A}_2^3\boldsymbol{\Psi}\boldsymbol{\Omega}_\tau\boldsymbol{\Psi}'\left(\mathcal{A}_2^3\right)'\boldsymbol{\phi}_{L_n}(\mathbf{z})$. This completes the proof of this theorem. ■

Proof of Theorem 3. The proof is only given for testing H_0(i) against H_1; the proof for testing H_0(ii) against H_1 can be derived similarly. When H_0(i) holds true, we first introduce some notation for our estimation under H_0(i) in parallel to that used in the proof of Theorem **1** or under H_1. To simplify notation, we let $\mathbf{z}_{i,n}$ include 1 in the model under the null and let $\boldsymbol{\chi}_{i,n} = (\mathbf{x}_{i,n}', \mathbf{z}_{i,n}', \mathbf{m}_{i,n}')'$.

Then, for a given value of ρ, the "profiled" estimator of $\boldsymbol{\theta}_{\tau,0}(\rho) = \left(\boldsymbol{\beta}_{\tau,0}(\rho)', \boldsymbol{\gamma}_{\tau,0}(\rho)', \boldsymbol{\delta}_{\tau,0}(\rho)'\right)'$ in Step 1 is given by

$$\widetilde{\boldsymbol{\theta}}_{\tau}(\rho) = \underset{\boldsymbol{\theta}_{\tau}(\rho)\in\boldsymbol{\Theta}}{\arg\min}\ \frac{1}{n}\sum_{i=1}^{n}\zeta_{\tau}\left(\dot{y}_{i,n}(\rho) - \boldsymbol{\chi}_{i,n}'\boldsymbol{\theta}_{\tau}(\rho)\right), \tag{A.13}$$

where $\boldsymbol{\theta}_{\tau}(\rho) = (\boldsymbol{\beta}_{\tau}(\rho)', \boldsymbol{\gamma}_{\tau}(\rho)', \boldsymbol{\delta}_{\tau}(\rho)')'$, $\widetilde{\boldsymbol{\theta}}_{\tau}(\rho) = \left(\widetilde{\boldsymbol{\beta}}_{\tau}(\rho)', \widetilde{\boldsymbol{\gamma}}_{\tau}(\rho)', \widetilde{\boldsymbol{\delta}}_{\tau}(\rho)'\right)'$, and $\boldsymbol{\theta}_{\tau,0}(\rho)$ is an interior point of a compact subset $\boldsymbol{\Theta}$.

Denote $\widetilde{\boldsymbol{\vartheta}}_{\tau}(\rho) = \sqrt{n}\left[\widetilde{\boldsymbol{\theta}}_{\tau}(\rho) - \boldsymbol{\theta}_{\tau}(\rho)\right]$, $Y_{i,n,0}^{*}(\rho) = y_{i,n} - \rho\sum_{j\neq i}w_{ij,n}y_{j,n} - \boldsymbol{\chi}_{i,n}'\boldsymbol{\theta}_{\tau,0}(\rho)$ and $Y_{i,n,0}(\rho) = Y_{i,n,0}^{*}(\rho) - n^{-1/2}\boldsymbol{\chi}_{i}'\widetilde{\boldsymbol{\vartheta}}_{\tau}(\rho)$. Again, solving for $\widetilde{\boldsymbol{\theta}}_{\tau}(\rho)$ is equivalent to finding $\widetilde{\boldsymbol{\vartheta}}_{\tau}(\rho)$ that minimizes

$$Q_{n,0}(\boldsymbol{\vartheta}_{\tau}(\rho)) = n^{-1}\sum_{i=1}^{n}\left[\zeta_{\tau}(Y_{i,n,0}(\rho)) - \zeta_{\tau}\left(Y_{i,n,0}^{*}(\rho)\right)\right]. \tag{A.14}$$

By definition, $RSC_{0,\tau} = \sum_{i=1}^{n}\zeta_{\tau}(\widetilde{u}_{i,n})$ and $RSC_{1,\tau} = \sum_{i=1}^{n}\zeta_{\tau}(\widehat{u}_{i,n})$ with

$$\begin{aligned}
\widetilde{u}_{i,n} &= y_{i,n} - \widetilde{\rho}_{\tau}\sum_{j\neq i}w_{ij,n}y_{j,n} - \mathbf{x}_{i,n}'\widetilde{\boldsymbol{\beta}}_{\tau} - \mathbf{z}_{i,n}'\widetilde{\boldsymbol{\delta}}_{\tau} = Y_{i,n,0}(\widetilde{\rho}_{\tau}) + \mathbf{m}_{i}'\widetilde{\boldsymbol{\gamma}}_{\tau}, \\
\widehat{u}_{i,n} &= y_{i,n} - \widehat{\rho}_{\tau}\sum_{j\neq i}w_{ij,n}y_{j,n} - \mathbf{x}_{i,n}'\widehat{\boldsymbol{\beta}}_{\tau} - \widehat{\alpha}_{\tau}(\mathbf{z}_{i,n}) = Y_{i,n}(\widehat{\rho}_{\tau}) + \mathbf{m}_{i}'\widehat{\boldsymbol{\gamma}}_{\tau},
\end{aligned}$$

are then calculated under the null and alternative hypotheses, respectively, where $Y_{i,n}(\rho)$ is defined in the same manner as in the proof of Theorem **1**.

Applying the same method as the one used to prove Theorem **1** , we obtain

$$\begin{aligned}
n^{-1}RSC_{0,\tau} &= n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y_{i,n,0}(\widetilde{\rho}_{\tau}) + \mathbf{m}_{i,n}'\widetilde{\boldsymbol{\gamma}}_{\tau}\right) \\
&= n^{-1}\sum_{i=1}^{n}\left[\zeta_{\tau}(Y_{i,n,0}(\widetilde{\rho}_{\tau})) - \zeta_{\tau}\left(Y_{i,n,0}^{*}(\widetilde{\rho}_{\tau})\right)\right] + n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y_{i,n,0}^{*}(\widetilde{\rho}_{\tau})\right) + \\
&\quad n^{-1}\sum_{i=1}^{n}\left[\zeta_{\tau}\left(Y_{i,n,0}(\widetilde{\rho}_{\tau}) + \mathbf{m}_{i,n}'\widetilde{\boldsymbol{\gamma}}_{\tau}\right) - \zeta_{\tau}(Y_{i,n,0}(\widetilde{\rho}_{\tau}))\right] \\
&= Q_{n,0}\left(\widetilde{\boldsymbol{\theta}}_{\tau}(\widetilde{\rho}_{\tau})\right) + n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y_{i,n,0}^{*}(\widetilde{\rho}_{\tau})\right) + n^{-1}\sum_{i=1}^{n}\left[\zeta_{\tau}\left(Y_{i,n,0}(\widetilde{\rho}_{\tau}) + \mathbf{m}_{i,n}'\widetilde{\boldsymbol{\gamma}}_{\tau}\right) - \zeta_{\tau}(Y_{i,n,0}(\widetilde{\rho}_{\tau}))\right] \\
&\approx -\frac{1}{2n^{2}}\left[\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\varphi_{\tau}(u_{i,n})\right]'\boldsymbol{\Sigma}_{\tau,0}^{-1}\left[\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\varphi_{\tau}(u_{i,n})\right] + n^{-1}\sum_{i=1}^{n}\zeta_{\tau}(u_{i,n}) + o_{p}\left(n^{-1}\right),
\end{aligned}$$

where $\boldsymbol{\Sigma}_{\tau,0} = \lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}f_{u_{i,n}}(0)\boldsymbol{\chi}_{i}\boldsymbol{\chi}_{i}'$. In addition, we obtain

$$\begin{aligned}
n^{-1}RSC_{1,\tau} &= n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y_{i,n}(\widehat{\rho}_{\tau}) + \mathbf{m}_{i,n}'\widehat{\boldsymbol{\gamma}}_{\tau}\right) \\
&= Q_{n}\left(\widehat{\boldsymbol{\theta}}_{\tau}(\widehat{\rho}_{\tau})\right) + n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y_{i,n}^{*}(\widehat{\rho}_{\tau})\right) + n^{-1}\sum_{i=1}^{n}\left[\zeta_{\tau}\left(Y_{i,n}(\widehat{\rho}_{\tau}) + \mathbf{m}_{i,n}'\widehat{\boldsymbol{\gamma}}_{\tau}\right) - \zeta_{\tau}(Y_{i,n}(\widehat{\rho}_{\tau}))\right] \\
&\approx -\frac{1}{2n^{2}}\left[\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_{\tau}(u_{i,n})\right]'\boldsymbol{\Sigma}_{\tau}^{-1}\left[\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_{\tau}(u_{i,n})\right] + n^{-1}\sum_{i=1}^{n}\zeta_{\tau}(u_{i,n}) + o_{p}\left(n^{-1}\right)
\end{aligned}$$

by the proof of Theorem **1**.

Therefore, under H_0, we obtain

$$RSC_{0,\tau} - RSC_{1,\tau} \approx \left(\mathcal{B}'_{n,1}\ \boldsymbol{\Sigma}_{\tau}^{-1}\mathcal{B}_{n,1} - \mathcal{B}'_{n,0}\ \boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}\right)/2,$$

where $\mathcal{B}_{n,0} = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\varphi_{\tau}(u_{i,n})$ and $\mathcal{B}_{n,1} = n^{-1/2}\sum_{i=1}^{n}\mathcal{X}_{i,n}\varphi_{\tau}(u_{i,n})$. By Seber (2008, Property 20.17), $\mathcal{B}'_{n,1}\boldsymbol{\Sigma}_{\tau}^{-1}\mathcal{B}_{n,1}$ can be rewritten as a linear combination of $d_x+d_m+L_n$ independent chi-squared random variables and $\mathcal{B}'_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}$ can be rewritten as a linear combination of $d_x+d_m+d_z$ independent chi-squared random variables. Since $n^{-1}RSC_{1,\tau} = n^{-1}\sum_{i=1}^{n}\zeta_{\tau}(u_{i,n}) + o_p(1) = n^{-1}\sum_{i=1}^{n}\mathbb{E}[\zeta_{\tau}(u_{i,n})] + o_p(1)$ by the LLN derived in Jenish & Prucha (2012), we obtain that

$$T_n = \frac{RSC_{0,\tau} - RSC_{1,\tau}}{RSC_{1,\tau}} \approx \frac{\left(\mathcal{B}'_{n,1}\boldsymbol{\Sigma}_{\tau}^{-1}\mathcal{B}_{n,1} - \mathcal{B}'_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}\right)/2}{\sum_{i=1}^{n}\mathbb{E}[\zeta_{\tau}(u_{i,n})]} = O_p\left(\frac{L_n}{n}\right).$$

Under H_1, following the proof of Theorem **1**, we can show that there exists parameter ρ_τ and $\boldsymbol{\theta}_\tau = \left(\boldsymbol{\beta}'_\tau, \boldsymbol{\gamma}'_\tau, \boldsymbol{\delta}'_\tau\right)'$ such that $\widetilde{\rho}_\tau - \rho_\tau = O_p\left(n^{-1/2}\right)$ and $\widetilde{\boldsymbol{\theta}}_\tau(\widetilde{\rho}_\tau) - \boldsymbol{\theta}_\tau = O_p\left(n^{-1/2}\right)$. Then, it follows that

$$n^{-1}RSC_{0,\tau} = n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y^*_{i,n,0}(\widetilde{\rho}_\tau) + \mathbf{m}'_{i,n}\widetilde{\boldsymbol{\gamma}}_\tau\right) \approx n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y^*_{i,n,0}(\rho_\tau)\right) = O_p(1),$$

where $Y^*_{i,n,0}(\rho_\tau) \equiv v_{i,n}(\rho_\tau) + \Delta_{i,n}$ and $\Delta_{i,n} = (\rho_{\tau,0} - \rho_\tau)\sum_{j=1}^{n} g_{ij,n}\left[\mathbf{x}'_{j,n}\boldsymbol{\beta}_{\tau,0} + \alpha_{\tau,0}(\mathbf{z}_{j,n})\right] + \mathbf{x}'_{i,n}\left(\boldsymbol{\beta}_{\tau,0} - \boldsymbol{\beta}_\tau\right) + \alpha_{\tau,0}(\mathbf{z}_{i,n}) - \mathbf{z}'_{i,n}\boldsymbol{\delta}_\tau - \mathbf{m}'_{i,n}\boldsymbol{\gamma}_\tau$. Because Lemma **1(ii)** shows that $\{v_{i,n}(\rho), \boldsymbol{l}(i) \in D_n\}$ is uniformly L_2-NED on $\{\varepsilon_{in}, \boldsymbol{l}(i) \in D_n\}$ with the NED coefficients of $\psi(s) = O(s^{-\varsigma})$, we apply the LLN of Jenish & Prucha (2012) and obtain $n^{-1}\sum_{i=1}^{n}\zeta_{\tau}\left(Y^*_{i,n,0}(\rho_\tau)\right) = n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\zeta_{\tau}\left(Y^*_{i,n,0}(\rho_\tau)\right)\right] + o_p(1)$. We therefore obtain, under H_1, that

$$T_n \xrightarrow{p} \frac{\varpi_0 - \varpi_1}{\varpi_1} > 0. \tag{A.15}$$

where we denote

$$\varpi_0 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\mathbb{E}\left[\zeta_{\tau}\left(Y^*_{i,n,0}(\rho_\tau)\right)\right] \text{ and } \varpi_1 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\mathbb{E}[\zeta_{\tau}(u_{i,n})]. \tag{A.16}$$

This completes the proof of this theorem. ■

Proof of Theorem 4. Again, the proof is only given for testing H_0(i) against H_1. We add a superscript "b" to the notation used in the proof of Theorem **3** for the bootstrap sample, the bootstrap estimator and the bootstrap test statistic. Closely following the proof of Theorem **3** and the proof method used in Feng et al. (2011, Supplemental Appendix), we have

$$\sum_{i=1}^{n}\mathbb{E}^b\left[\zeta_{\tau}\left(u^b_{i,n}\right)\right]T^b_n \approx \left(\mathcal{B}^{b\prime}_{n,1}\boldsymbol{\Sigma}_{\tau}^{-1}\mathcal{B}^b_{n,1} - \mathcal{B}^{b\prime}_{n,0}\boldsymbol{\Sigma}_{\tau,0}^{-1}\mathcal{B}^b_{n,0}\right)/2.$$

We also have that $\mathbb{E}^b\left[\zeta_{\tau}\left(u^b_{i,n}\right)\right] = 2\tau(1-\tau)|\widetilde{u}_{i,n}|$, where we let $\mathbb{E}^b[\cdot] \equiv \mathbb{E}[\cdot|\{(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, y_{i,n})\}]$. In addition, conditional on $\{(\mathbf{x}_{i,n}, \mathbf{z}_{i,n}, y_{i,n})\}$, $\mathcal{B}^b_{n,0} = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\varphi_{\tau}\left(u^b_{i,n}\right)$ converges to a normal distribution with zero mean and variance equal to $n^{-1}\tau(1-\tau)\sum_{i=1}^{n}\boldsymbol{\chi}_{i,n}\boldsymbol{\chi}'_{i,n}$, which is the same asymptotic distribution as that of $\mathcal{B}_{n,0}$ under H_0, while $\mathcal{B}_{n,1}^b = n^{-1/2}\sum_{i=1}^n \mathcal{X}_{i,n}\varphi_\tau\left(u_{i,n}^b\right)$ converges to a normal distribution with zero mean and variance equal to $n^{-1}\tau(1-\tau)\sum_{i=1}^n \mathcal{X}_{i,n}\mathcal{X}_{i,n}'$, which is the same asymptotic distribution as that of $\mathcal{B}_{n,1}$ under H_0. Therefore, we obtain $T_n^b = O_p(L_n/n)$ under both H_0 and H_1, and

$$J_n^b = 2\tau(1-\tau)\left(n^{-1}\sum_{i=1}^n |\widetilde{u}_{i,n}|\right) nT_n^b \approx \left(\mathcal{B}_{n,1}^{b\prime}\mathbf{\Sigma}_\tau^{-1}\mathcal{B}_{n,1}^b - \mathcal{B}_{n,0}^{b\prime}\mathbf{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}^b\right)/2$$

has an asymptotic conditional distribution same as the asymptotic distribution of $\left(\mathcal{B}_{n,1}'\mathbf{\Sigma}_\tau^{-1}\mathcal{B}_{n,1} - \mathcal{B}_{n,0}'\mathbf{\Sigma}_{\tau,0}^{-1}\mathcal{B}_{n,0}\right)/2$ under H_0.

Bickel & Freedman (1981) show that a convergence in the Mallows distance is equivalent to the weak convergence, and Pólya's theorem extends the weak convergence to the strong convergence (Serfling, 2002, p.18). Also, under H_1, we have $J_n = \left(n^{-1}\sum_{i=1}^n \zeta_\tau(\widehat{u}_{i,n})\right) nT_n = O_p(n)$, which implies that $\Pr[J_n > c_a] \to 1$. Therefore, the bootstrap statistic is consistent. This completes the proof of this theorem. ■

References

Anselin, L. & Lozano-Gracia, N. (2009). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, *34*, 5–34.

Bajari, P., Fruehwirth, J. C., Kim, K. I., & Timmins, C. (2012). A rational expectations approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution. *American Economic Review*, *102*, 1898–1926.

Baltagi, B. & Li, D. (2001). LM tests for functional form and spatial correlation. *International Regional Science Review*, *24*, 194–225.

Bickel, P. J. & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, *9*, 1196–1217.

Brasington, D. M. & Hite, D. (2005). Demand for environmental quality: A spatial hedonic analysis. *Regional Science and Urban Economics*, *35*, 57–82.

Cai, Z., Das, M., Xiong, H., & Wu, X. (2006). Functional coefficient instrumental variables models. *Journal of Econometrics*, *133*(1), 207–241.

Cai, Z. & Li, Q. (2008). Nonparametric estimation of varying coefficient dynamic panel data models. *Econometric Theory*, *24*(5), 1321–1342.

Cai, Z. & Xu, X. (2008). Nonparametric quantile estimation for dynamic smooth coefficient models. *Journal of the American Statistical Association*, *103*, 1595–1608.

Chay, K. Y. & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, *113*, 376–424.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of Econometrics*, volume 6B. North Holland.

Chen, X. & Pouzo, D. (2009). Efficient estimation of semiparametric conditional moment models with possibly nonsmooth residuals. *Journal of Econometrics*, *152*, 46–60.

Chen, X. & Pouzo, D. (2012). Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals. *Econometrica*, *80*, 277–321.

Chen, X. & Pouzo, D. (2015). Sieve Wald and QLP inferences on semi/nonparametric conditional moment models. *Econometrica*, *83*, 1013–1079.

Chernozhukov, V. & Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, *132*, 491–525.

Cohen, J. P. & Coughlin, C. C. (2008). Spatial hedonic models of airport noise. *Journal of Regional Science*, *48*, 859–878.

Fan, J., Hu, T., & Truong, Y. K. (1994). Robust nonparametric function estimation. *Scandinavian Journal of Statistics*, *21*, 433–446.

Fan, J., Zhang, C., & Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Annals of Statistics*, *29*, 153–193.

Feng, X., He, X., & Hu, J. (2011). Wild bootstrap for quantile regression. *Biometrika*, *98*, 995–999.

Gagliardini, P. & Scaillet, O. (2012). Nonparametric instrumental variable estimation of structural quantile effects. *Econometrica*, *80*, 1533–1562.

Gawande, K. & Jenkins-Smith, H. (2001). Nuclear waste trasport and residential property values: Estimating the effects of perceived risks. *Journal of Environmental Economics and Management*, *42*, 207–233.

Gayer, T. (2000). Neighborhood demographics and the distribution of hazardous waste risks: An instrumental variables estimation. *Journal of Regulatory Economics*, *17*, 131–155.

Greenstone, M. & Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the Superfund program. *Quarterly Journal of Economics*, *123*, 951–1003.

Harrison, D. & Rubenfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management*, *5*, 81–102.

He, X. & Liang, H. (2000). Quantile regression estimates for a class of linear and partially linear errors-in-variables models. *Statistica Sinica*, *10*, 129–140.

He, X. & Shi, P. (1996). Bivariate tensor-product B-splines in a partly linear model. *Journal of Multivariate Analysis*, *58*, 162–181.

Hite, D., Chern, W., Hitzhusen, F., & Randall, A. (2001). Property-value impacts of an environmental disamenity: The case of landfills. *Journal of Real Estate Finance and Economics*, *22*, 185–202.

Honda, T. (2004). Quantile regression in varying coefficient models. *Journal of Statistical Planning and Inferences*, *121*, 113–125.

Horowitz, J. L. & Lee, S. (2007). Nonparametric instrumental variable estimation of a quantile regression model. *Econometrica*, *75*, 1191–1208.

Jaeger, W. (2006). The hidden costs of relocating sand and gravel mines. *Resources Policy*, *31*, 146–164.

Jenish, N. (2016). Spatial semiparametric model with endogenous regressors. *Econometric Theory*, *32*, 714–739.

Jenish, N. & Prucha, I. (2009). Central limit theorems and uniform law of large numbers for arrays of random fields. *Journal of Econometrics*, *150*, 86–98.

Jenish, N. & Prucha, I. (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, *170*, 178–190.

Kelejian, H. H. & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive models with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, *17*, 99–121.

Kelejian, H. H. & Prucha, I. R. (1999). A generalized moment estimator for the autoregressive parameter in a spatial model. *International Economic Review*, *40*, 509–533.

Kelejian, H. H. & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, *157*, 53–67.

Kim, M.-O. (2007). Quantile regression with varying coefficients. *Annals of Statistics*, *35*, 92–108.

Koenker, R. & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*, 1296–1310.

Kohlhase, J. E. (1991). The impact of toxic waste sites on housing values. *Journal of Urban Economics*, *30*, 1–26.

Lee, S. (2003). Efficient semiparametric estimation of a partially linear quantile regression model. *Econometric Theory*, *19*, 1–31.

Lee, T.-H. & Ullah, A. (2003). Nonparametric bootstrap specification testing in econometric models. In D. E. Giles (Ed.), *Computer-Aided Econometrics* (pp. 451–477). NY: CRC Press.

Leggett, C. G. & Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management*, *39*, 121–144.

LeSage, J. & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Boca Raton: Taylor & Francis CRC Press.

Li, Q. & Stengos, T. (1996). Semiparametric estimation of partially linear panel data models. *Journal of Econometrics*, *71*, 389–397.

Li, Q. & Ullah, A. (1998). Estimating partially linear models with one-way error components. *Econometric Reviews*, *17*, 145–166.

Liao, W.-C. & Wang, X. (2012). Hedonic house prices and spatial quantile regression. *Journal of Housing Economics*, *21*, 16–27.

Malikov, E. & Sun, Y. (2017). Semiparametric estimation and testing of smooth coefficient spatial autoregressive models. *Journal of Econometrics*, *199*, 12–34.

Nelson, A. C., Genereux, J., & Genereux, M. (1992). Price effects of landfills on house values. *Land Economics*, *68*, 359–365.

Newey, W. K. & McFadden, D. (1994). Large sample estimation and hypothesis testing. In R. Engle & D. McFadden (Eds.), *Handbook of Econometrics*, volume IV chapter 36, (pp. 2111–2245). Elsevier Science B.V.

Ohio Department of Natural Resources (2010). *2009 Report on Ohio Mineral Industries: An Annual Summary of the State's Economic Geology*. Division of Geological Survey. Compiled by Mark E. Wolfe.

Ohio Department of Natural Resources (2011). *2010 Report on Ohio Mineral Industries: An Annual Summary of the State's Economic Geology*. Division of Geological Survey. Compiled by Mark E. Wolfe.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory*, *7*, 186–199.

Qu, X. & Lee, L.-f. (2015). Estimating a spatial autoregressive model with an endogeneous spatial weight matrix. *Journal of Econometrics*, *184*, 209–232.

Reichert, A. K., Small, M., & Mohanty, S. (1992). The impact of landfills on residential property values. *Journal of Real Estate Research*, *7*, 297–314.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*, 34–55.

Ruppert, D. & Carroll, R. J. (1980). Trimmed least squares estimation in the linear model. *Journal of the American Statistical Association*, *75*, 828–838.

Seber, G. A. F. (2008). *A Matrix Handbook for Statisticians*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc.

Small, K. A. & Steimetz, S. S. C. (2012). Spatial hedonics and the willingness to pay for residential amenities. *Journal of Regional Science*, *52*, 635–647.

Su, L. (2012). Semiparametric GMM estimation of spatial autoregressive models. *Journal of Econometrics*, *167*, 543–560.

Su, L. & Hoshino, T. (2016). Sieve instrumental variable quantile regression estimation of functional coefficient models. *Journal of Econometrics*, *191*, 231–254.

Su, L. & Jin, S. (2010). Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *Journal of Econometrics*, *157*, 18–33.

Su, L. & Yang, Z. (2011). Instrumental variable quantile estimation of spatial autoregressive models. Working Paper, Singapore Management University.

Sun, Y. (2006). A consistent nonparametric equality test of conditional quantile functions. *Econometric Theory*, *22*, 614–632.

Sun, Y., Hongjia, Y., Zhang, W., & Lu, Z. (2014). A semiparametric spatial dynamic model. *Annals of Statistics*, *42*, 700–727.

Sun, Y. & Malikov, E. (2018). Estimation and inference in functional-coefficient spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*. forthcoming.

Ullah, A. (1985). Specification analysis of econometric models. *Journal of Quantitative Economics*, *1*, 187–209.

Yu, K. & Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association*, *93*, 228–237.

Zabel, J. E. & Guignet, D. (2012). A hedonic analysis of the impact of LUST sites on house prices. *Resource and Energy Economics*, *34*, 549–564.

Zhang, Z. (2013). A pairwise difference estimator for partially linear spatial autoregressive models. *Spatial Economic Analysis*, *8*, 176–194.