Online Appendix:
The Housing Market Impacts of Shale Gas Development

Lucija Muehlenbachs, Elisheba Spiller, and Christopher Timmins

August 2015

A Sample Cuts

We start with 731,169 unique observations of sales that have information on the location of the property. After excluding properties without a listed price, a price in the top or bottom 1 percent of all prices, and properties sold more than once in a single year, we are left with 626,637 sales observations. Of these, there are 604,074 properties designated as a single family residence, rural home site, duplex, or townhouse; our main specifications only include these properties in order to estimate the impact on (likely) owner-occupied homes, rather than properties that are more likely transient or rented.[1] Furthermore, we want to include in our main specification only homes that were sold from one person to another (i.e., excluding made-to-order homes), thus we drop approximately 30,203 properties that were sold in the year built.[2] After eliminating new homes, of the remaining 573,871 sales, 229,946 are repeat sales—a necessary condition for including property fixed effects. For specifications that instead rely on observed housing attributes (specifically, our Linden and Rockoff-type figures), not all properties report a full slate of housing characteristics; out of our 573,871 sale sample, only 379,649 have information on all property characteristics.

B Hedonic Method

Rosen (1974) established the connection between individual preferences and the hedonic price function, allowing the researcher to interpret the hedonic gradient as the marginal willingness to pay for an incremental change in a non-marketed house or neighborhood attribute. In the context of our application, $P(W)$ represents the hedonic price relationship describing how prices vary with exposure to increasing numbers of wells, *ceteris paribus*. Rosen describes how the hedonic price function

[1] Though CoreLogic provides an indicator for whether the property is owner-occupied, this variable is not consistently reported by all counties. We exclude properties listed as a hotel, motel, residence hall, or transient lodging.

[2] Results are similar if these homes are included. We return to the question of new home construction in response to shale gas development in Appendix Section G.

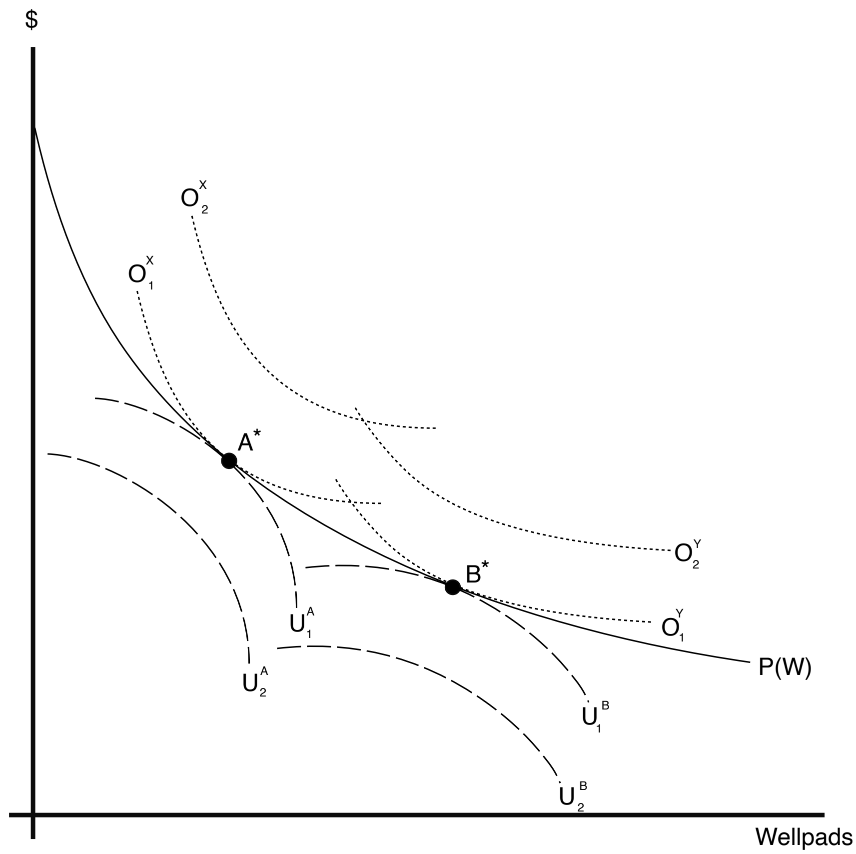

Figure B.1: Formation of the Hedonic Price Function

is formed by the equilibrium of buyers and sellers sorting to one another in the marketplace. In Figure B.1, buyers A and B are represented by indifference curves (U_1^A, U_1^B, U_2^A, U_2^B); each represents combinations of price and shale gas well exposure that yield a constant level of utility. Sellers X and Y are described by offer curves (O_1^X, O_1^Y, O_2^X, O_2^Y), each of which represents combinations of price and well exposure that yield a constant level of profit. The hedonic price function is formed by the envelope of these indifference and offer curves.

Individuals choose a house that maximizes utility. For individual A, who neither likes paying a lot for a house nor (for the purposes of this discussion) wants exposure to shale gas wells, this is accomplished by reaching the indifference curve lying farthest to the southwest. Considering the constraint formed by the hedonic price function, utility is maximized at point A^*, where that individual achieves utility U_1^A. Individual B similarly maximizes utility at B^*. The fundamental insight of the hedonic method is that, at A^* and B^*, the slope of the price function is equal to the slope of each individual's indifference curve at that point. That slope describes the individual's willingness to give up consumption of other goods in exchange for a marginal reduction in exposure to nearby wells. This is how the literature typically defines marginal willingness to pay (MWTP); we will do the same.[3]

Of course, the value of MWTP defined by the slope of the price function at the level of well exposure chosen by the individual represents just one point on the individual's indifference curve. If we were to trace out each individual's MWTP at each point on a particular indifference curve, we would end up with functions for each individual like those shown in Figure B.2.

With cross-sectional data, the hedonic gradient (i.e., the slope of the hedonic price function) therefore only identifies one point on each MWTP function. This is the crux of the identification problems detailed by Brown and Rosen (1982) and Mendelsohn (1985). Endogeneity problems also arise in the effort to econometrically recover these functions; for a discussion, see Bartik (1987) and Epple (1987). More recent literature dealing with the recovery of MWTP functions includes Ekeland et al. (2004), Bajari and Benkard (2005), Heckman et al. (2010), and Bishop and Timmins (2012).

With few exceptions, the applied hedonic literature has not estimated heterogeneous MWTP functions, but has instead relied on a strong assumption to simplify the problem—in particular, that preferences are homogenous and are therefore represented by the hedonic price function itself. Using price levels as the dependent variable (so that the hedonic gradient is a horizontal line that represents the MWTP function for all individuals) yields the simple estimate of MWTP in Figure B.3, and avoids the difficulties associated with recovering estimates of MWTP discussed above (using log prices, these become non-linear functions, but also allow us to recover a simple estimate of MWTP without having to estimate the MWTP function). This has allowed attention to be focused instead on recovering unbiased estimates of the hedonic price function. This literature is vast and includes applications dealing with air quality (Chay and Greenstone, 2005; Bajari et al., 2010; Bui and Mayer, 2003; Harrison Jr and Rubinfeld, 1978; Ridker and Henning, 1967),

[3] Other measures of value used in the literature include compensating and equivalent variations in income. CV or EV can be calculated both in a partial equilibrium context, where individuals' housing choices and equilibrium prices are not updated, and in a general equilibrium context, where they are updated to reflect re-optimization and subsequent market re-equilibration.

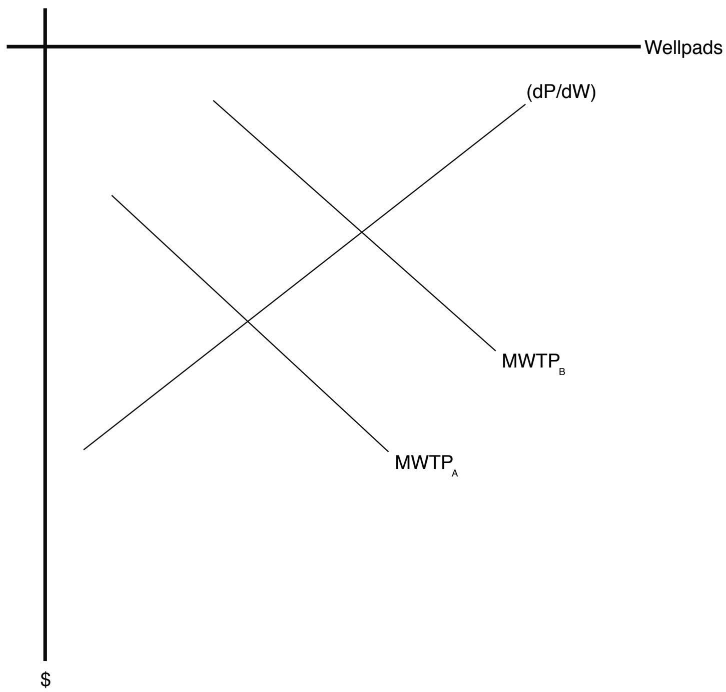

Figure B.2: Marginal Willingness to Pay

water quality (Walsh et al., 2011; Poor et al., 2007; Leggett and Bockstael, 2000), school quality (Black, 1999), crime (Linden and Rockoff, 2008; Pope, 2008b), and airport noise (Andersson et al., 2010; Pope, 2008a). Our application is most similar in spirit to papers that have examined locally undesirable land uses (LULUs): Superfund sites (Greenberg and Hughes, 1992; Kiel and Williams, 2007; Greenstone and Gallagher, 2008; Gamper-Rabindran and Timmins, 2013), brownfield redevelopment (Haninger et al., 2012; Linn, 2013), commercial hog farms (Palmquist et al., 1997), underground storage tanks (Zabel and Guignet, 2012), cancer clusters (Davis, 2004), and electric power plants (Davis, 2011). Our estimation strategy described above draws upon insights from many of these papers. We follow the literature and specify a log-linear hedonic price function. As in these other papers, the smaller the change in the (dis)amenity, the better this function will approximate the true partial equilibrium welfare effect.

Of particular importance for our analysis is the discussion in Kuminoff and Pope

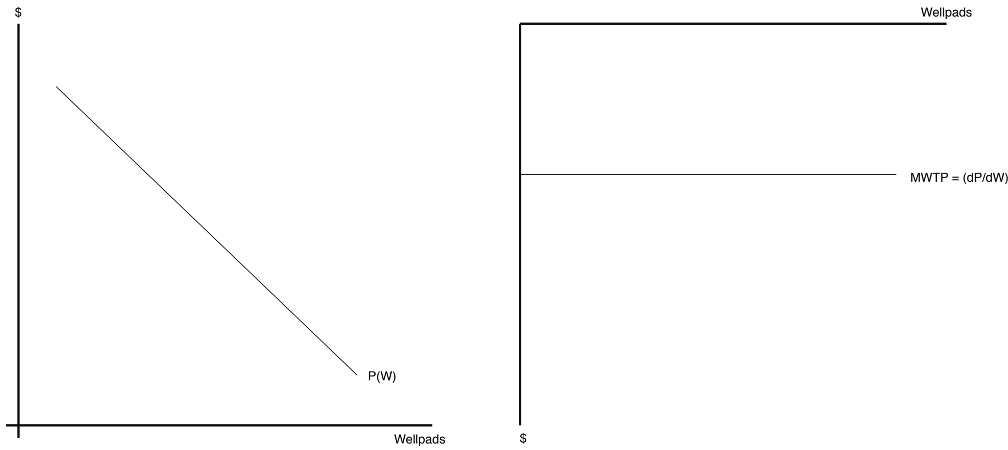

Figure B.3: Marginal Willingness to Pay—Simplification

(2014). They highlight the fact that the change in price over time (which allows for the use of differencing strategies to control for time-invariant unobservables) will only yield a measure of the willingness to pay for the corresponding change in the attribute being considered under a strong set of assumptions. These assumptions include those described above (i.e., linear hedonic price function, common MWTP function). In addition, the hedonic price function must not move over the time period accompanying the change in the attribute. If it does, as in Figure B.4, the change in the price (δP) accompanying the change in well pad exposure (δW) may provide a poor approximation of the slopes of either of the hedonic price functions. [4] In the right panel of Figure B.4, $\left|\frac{\Delta P}{\Delta W}\right|$ (i.e., the dashed line) $<$ $\left|\frac{dP}{dW}\right|$ (i.e., the solid line).

Determining whether or not the hedonic price function has moved over time is difficult; in particular, it requires having some way of recovering an unbiased estimate of the hedonic price function without exploiting time variation. As a check on our DDD results, we provide an alternative strategy for recovering the impact of groundwater contamination risk (double-difference nearest neighbor matching)

[4] Even if the hedonic price function remains stationary over time, the change in price accompanying the change in an amenity will not accurately describe the slope of the price function if that function is non-linear. The problem becomes more severe the larger the change in the amenity being considered. Many papers in the hedonics literature described above consider non-marginal changes. Our estimation looks at small, marginal changes in the number of wells adjacent to a property.

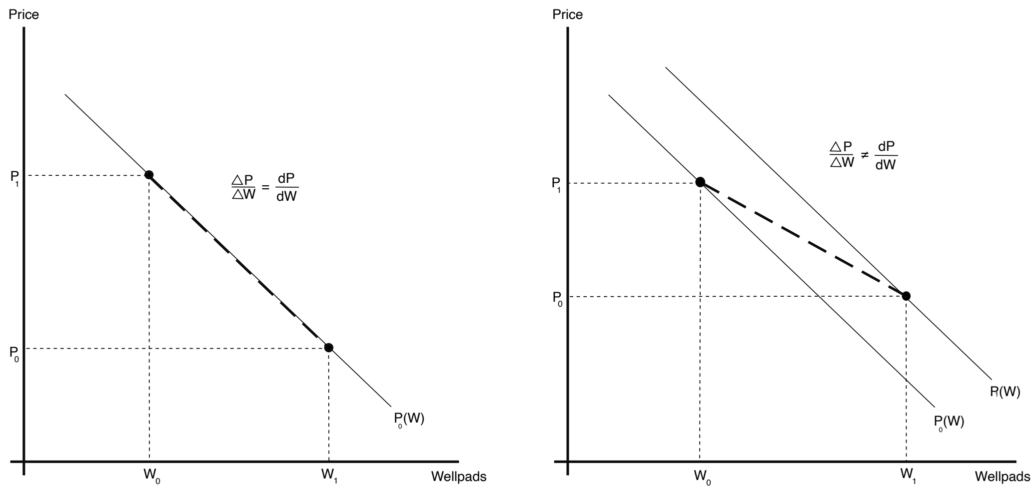

Figure B.4: Time-Varying $P(W)$

that avoids using time variation. In the following sub-section of the appendix, we also provide an indication of how much of a problem shifting gradients present for our double- and triple-difference strategies by looking at the extent to which neighborhood sociodemographics change because of fracking. If they change a lot, preferences of the local population will likely be altered as well, and caution would be advised when interpreting our results as measures of welfare rather than simple capitalization effects. We note here, however, that the changes we find attributable to shale gas development are quite small.

C Groundwater Contamination Risk and Adjacency Impacts Beyond k=2km

In this subsection, we extend Tables 2 and 3 to include regressors of the counts of wells at 2.5km and 3km. At farther distances, an additional well has a smaller effect on PWSA properties and no impact on GW properties (the last four columns of Table 2). We see that the adjacency impacts remain the same or decrease with radii larger than 2km in the case of the different types of adjacent wells (last two columns of Table 3). That an additional well has a smaller impact the farther from well, could be driven by farther wells having a smaller impact, but also by non-linear effects because there are more wells found in larger radii. We cannot rule out that the impact that well pads have on properties is non-linear in the number of well pads. We do not have enough variation in the number of well pads to reliably estimate non-linear effects. Restricting the sample to only those properties that eventually have at most one well within 2km (not shown), we do not have significance with a radius of 1km (possibly due to the small sample size) and find much larger impacts with radii of 1.5km and 2km than in our main table.

Table C.1: Log Sale Price Well Pads

	$K \leq 1$km		$K \leq 1.5$km		$K \leq 2$km		$K \leq 2.5$km		$K \leq 3$km	
	Full	Bound.	Full	Bound.	Full	Bound.	Full	Bound.	Full	Bound.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pads in Kkm	.028	.026	.029**	.034*	.016**	.018*	.012**	.014*	.011***	9.9e-03*
	(.025)	(.035)	(.014)	(.02)	(6.9e-03)	(.01)	(4.9e-03)	(7.2e-03)	(3.4e-03)	(5.6e-03)
(Pads in Kkm)*GW	-.062	-.165**	-.042*	-.099***	-.023	-.013	-.013	-7.0e-03	-.01	-.012
	(.046)	(.072)	(.025)	(.036)	(.02)	(.052)	(.014)	(.029)	(9.7e-03)	(.023)
Pads in 20km	-7.8e-04***	-8.1e-04	-8.3e-04***	-9.3e-04*	-8.4e-04***	-9.4e-04*	-8.7e-04***	-1.0e-03*	-9.0e-04***	-1.0e-03*
	(3.0e-04)	(5.3e-04)	(3.0e-04)	(5.5e-04)	(3.0e-04)	(5.6e-04)	(3.0e-04)	(5.7e-04)	(3.0e-04)	(5.7e-04)
(Pads in 20km)*GW	6.6e-04	2.0e-03***	7.0e-04	2.0e-03***	7.1e-04	1.7e-03**	7.0e-04	1.7e-03**	7.1e-04	1.9e-03***
	(4.7e-04)	(7.0e-04)	(4.9e-04)	(6.8e-04)	(5.2e-04)	(6.8e-04)	(4.9e-04)	(6.7e-04)	(5.0e-04)	(6.8e-04)
Property Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	229,946	66,327	229,946	66,327	229,946	66,327	229,946	66,327	229,946	66,327
p-value ($\alpha_3 + \alpha_4 = 0$)	.414	.051	.544	.090	.740	.919	.935	.817	.950	.928
Avg. Pads in Kkm	.003	.006	.009	.015	.018	.031	.031	.055	.048	.081
Avg. Pads in 20km	4.725	5.108	4.725	5.108	4.725	5.108	4.725	5.108	4.725	5.108

Notes: This table extends the first panel of Table 2 to radii beyond K=2km. Dependent variable is log sale price and each column represents a separate regression. The independent variables in the regressions vary by the size of the radius Kkm around each property, used to count the number of adjacent well pads present before the sale. The boundary sample restricts the full sample to include only properties in a narrow band around the border of the public water service areas. Robust standard errors are clustered by census tract. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table C.2: Adjacency Effects

	K=1km	K=1.5km	K=2km	K=2.5km	K=3km
	(1)	(2)	(3)	(4)	(5)
	ln(price)	ln(price)	ln(price)	ln(price)	ln(price)
A. Log Sale Price on Well Pads in View					
Visible Well Pads in Kkm	1.1e-03	-.019	.019	.018	6.0e-03
	(.072)	(.058)	(.035)	(.02)	(.012)
Not-Visible Well Pads in Kkm	.03	.036***	.015**	.011**	.011***
	(.028)	(.013)	(6.5e-03)	(4.6e-03)	(3.4e-03)
Pads in 20km	-6.0e-04*	-6.4e-04*	-6.5e-04*	-6.8e-04**	-7.1e-04**
	(3.3e-04)	(3.3e-04)	(3.3e-04)	(3.4e-04)	(3.4e-04)
B. Log Sale Price on Productive Wells					
Unproductive Pads in Kkm	-.052	-.043	-.054*	-.03	6.7e-03
	(.077)	(.035)	(.03)	(.022)	(.02)
Producing Pads in Kkm	.044**	.038***	.02***	.014***	.011***
	(.02)	(.013)	(5.8e-03)	(4.5e-03)	(3.3e-03)
Pads in 20km	-6.0e-04*	-6.4e-04*	-6.3e-04*	-6.4e-04*	-7.0e-04**
	(3.3e-04)	(3.3e-04)	(3.3e-04)	(3.4e-04)	(3.4e-04)
C. Log Sale Price on Timing of Wellbores					
Old Bores (drilled $>$ 365 days) in Kkm	.021	.023**	.011**	.011***	9.6e-03***
	(.018)	(9.8e-03)	(4.4e-03)	(3.3e-03)	(2.6e-03)
New Bores (drilled $\leq$ 365 days) in Kkm	-4.4e-03	-9.7e-03	-3.3e-04	-6.0e-03	-1.9e-03
	(.029)	(.013)	(8.0e-03)	(6.3e-03)	(5.2e-03)
Old Undrilled Permits ($>$ 365 days) in Kkm	.055**	.022	.011	9.8e-03	6.4e-03
	(.025)	(.014)	(.012)	(8.9e-03)	(7.0e-03)
New Undrilled Permits ($\leq$ 365 days) in Kkm	.04*	7.2e-03	7.2e-03	2.0e-03	-1.3e-03
	(.023)	(.014)	(7.9e-03)	(5.9e-03)	(4.9e-03)
Pads in 20km	-6.0e-04*	-6.2e-04*	-6.3e-04*	-6.4e-04*	-6.8e-04**
	(3.3e-04)	(3.3e-04)	(3.3e-04)	(3.4e-04)	(3.4e-04)
Property Effects	Yes	Yes	Yes	Yes	Yes
County-Year Effects	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes
n	212,207	212,207	212,207	212,207	212,207

Notes: This table extends Table 3 to radii beyond K=2km. Dependent variable is log sale price. Each panel has three separate regressions, one per column. Regressors are the count of wells (or annual natural gas production) within Kkm, depending on the column. The sample used includes only properties that are in piped water service areas. Robust standard errors are clustered by census tract. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

D Event Study using the Timing of Drilling

In this subsection we create event-study graphs similar to Greenstone and Hanna (2014) for four types of properties, in which the event examined is the drilling of the first well. We create indicators for each of the years before and after the first well is drilled for properties that are adjacent to a well (within 2km) and properties that are nearby but not adjacent (within 2-10km). We separate the sample into subsamples depending on whether the property is at some point in time adjacent ("treated") or only nearby ("control") and whether the property has access to piped water or is dependent on groundwater. For each subsample, we run separate regressions of logged property values on the dummies indicating how many years (before or after) there are between the sale and drill date. Because the timing of the drilling varies across different properties we can identify year fixed effects as well the coefficients on the dummies and so we can control for year effects and quarter effects. Figures D.1 and D.2 plot the coefficients on each of the dummies (in which the omitted category is the dummy indicated the property was sold seven years before the first well). These figures are useful to demonstrate that there are no differential trends between the treatment and control groups in the years prior to the drilling. In the pre-period, the coefficients on the years prior to drilling are statistically insignificant in all subsamples. In the years after drilling, we have statistical significance on GW properties in the treatment group in the second year after drilling. Similar to Table 3, Panel C, these estimates suggest that there is a delay in the impacts. This implies that our main estimation, by considering the effect on adjacent GW houses in all years after drilling (including the year directly after drilling), may be understating the size of the total effect. However, similar to the Linden and Rockoff (2008) approach, this test is using a particular sample (i.e., only looking at homes exposed to a well pad within 10km) which is not necessarily representative of all homes affected by proximate wells. Therefore, our preferred specification comes from the estimation on the full set of well pads, the boundary sample, and the triple difference.

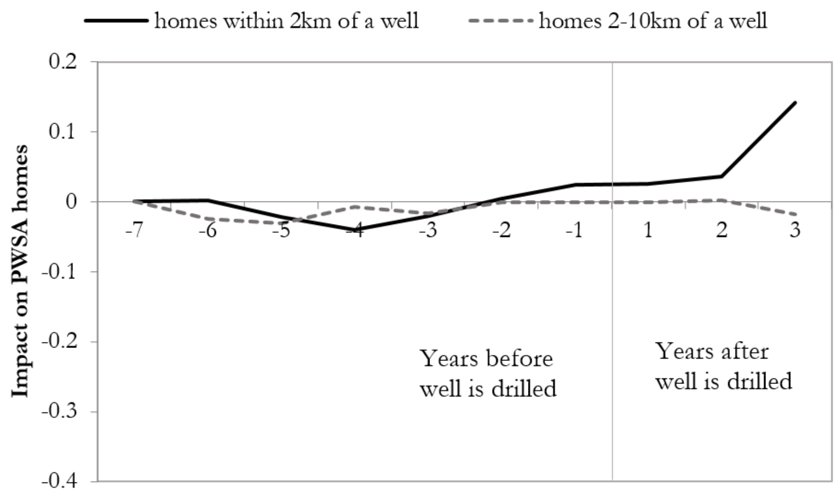

Figure D.1: Years Since First Well Drilled for Properties with Access to Piped Water

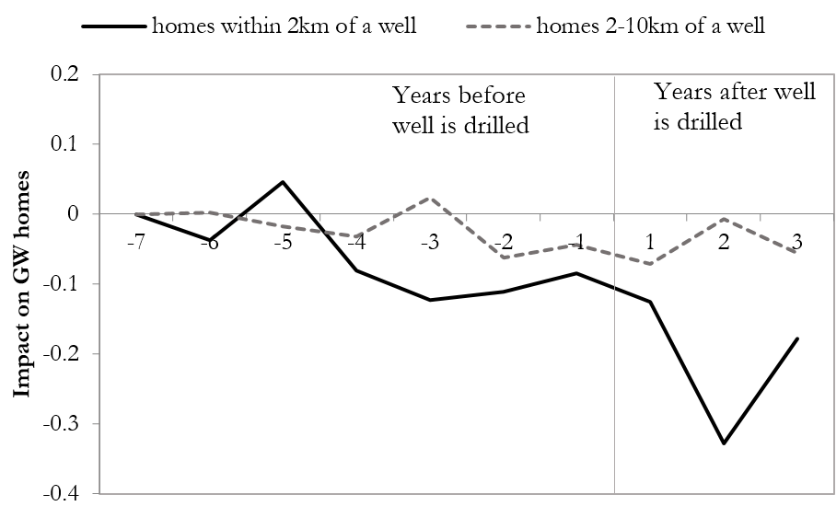

Figure D.2: Years Since First Well Drilled for Properties Dependent on Groundwater

E Effects on Sociodemographics

In this subsection, we examine the effect of shale gas development on sociodemographic attributes at the census-tract level. As described in Section B, if the hedonic price function moves over time, the change in price accompanying a change in exposure to shale gas may provide a poor approximation of the slope of the hedonic price function. Kuminoff and Pope (2014) discuss a number of conditions that must hold in order for this not to be a concern. One important requirement is that the preferences of local residents for exposure to wells do not change over time. If preferences are a function of residents' attributes, a simple check can be performed by examining how tract-level sociodemographics change with changes in exposure. To examine how sociodemographics changed over time, we compare 2000 and 2012 using census tract information on neighborhood attributes compiled by SimplyMap, a national data mapping software tool.[5] SimplyMap combines information from decennial censuses, the American Community Survey Public Use Microdata Samples, the Annual Demographic Survey, Current Population Reports, numerous special Census reports, and information from the US Postal Service to create estimates for key sociodemographic variables at the census tract level. Table E.1 describes the results of this analysis. In particular, we regress the change in 33 tract-level attributes, X, over the period 2000 to 2012 on the change in the number of cumulative wellbores within 20km of the centroid of the census tract in 2012.[6]

$$(X_{i,2012} - X_{i,2000}) = \rho\, bores20_{i,2012} + \epsilon_i$$

The first column reports the variable name, and the second column reports the mean of that variable in 2012. The third column reports the coefficient on wellbores, ρ, and the fourth column reports the percent change in the variable in question over the period 2000 to 2012 attributable to the average change in the number of wells in the corresponding vicinity of each census tract.

Out of the 33 variables that we consider, 23 have statistically significant wellbore effects. While statistical significance may be a cause for concern, very few of these effects are *economically* significant. In particular, considering the actual change in well exposure in each census tract over this period, the average of the resulting changes in tract attributes was no larger than 1 percent for any variable. Changes in neighborhood composition induced by shale gas development are, therefore, quite small. While this is not sufficient to rule out shifts in the hedonic price function over time, it is evidence in favor of a MWTP, as opposed to a simple capitalization

[5] http://geographicresearch.com/simplymap/. Access through the Duke University Library.

[6] Recall that cumulative wellbores is everywhere equal to zero in 2000.

effect, interpretation of our DDD results.

Table E.1: Change in Sociodemographic Characteristics, 2000-2012

Variable	Mean in 2012	Coefficient on Wellbores	Average % Δ from Wells
Household Income per Capita	30,080.30	-2.45E0	-0.154
Household Median Vehicles	1.803	1.30E-4***	0.071
Median Age	39.09	5.83E-3***	0.156
Median Age (Female)	40.294	5.19E-3***	0.135
Median Age (Male)	37.706	6.87E-3***	0.189
Population	3,964.24	-6.05E-1***	-0.291
% Asian	0.059	-6.25E-5***	-0.009
% Associate Degree	0.055	3.10E-5***	0.000
% Bachelor's Degree	0.122	-2.24E-6	0.000
% Black	0.155	-6.62E-6	0.000
% Family	0.784	-1.59E-5	0.000
% Female	0.515	-2.39E-5***	0.000
% High School	0.211	2.74E-5***	0.000
% Hispanic	0.131	-9.98E-5***	-0.004
% In Group Quarters	0.034	6.69E-6	0.001
% Less Than High School	0.093	-3.46E-5***	0.000
% Male	0.485	2.39E-5***	0.000
% Married, Female	0.202	-2.91E-5***	0.000
% Married, Male	0.204	-3.52E-5***	0.000
% Non-Family	0.182	9.22E-6	0.000
% Occupation, Construction	0.034	-1.05E-5**	0.000
% Occupation, Farming	0.002	-1.17E-6	0.000
% Occupation, Management	0.068	-1.07E-5	0.000
% Occupation, Production	0.054	-9.87E-6*	0.000
% Occupation, Professional	0.107	8.36E-7	0.000
% Occupation, Sales and Office	0.111	1.11E-5	0.000
% Occupation, Service	0.092	-1.81E-5**	0.000
% Other Race	0.052	5.56E-5***	0.013
% Some College	0.115	2.43E-5***	0.000
% Speaks English	0.728	1.16E-4***	0.000
% Urban	0.835	-9.92E-6***	0.000
% White	0.701	7.68E-5***	0.000
% White, Non-Hispanic	0.643	1.33E-4***	0.000

Note: % Δ from Wells is calculated as the average across census tracts of (Δ Wellbores*Coefficient on Wellbores)/(Mean in 2012)*100.

F Effects on Likelihood of Transaction

Here we investigate whether shale gas development within 20km affects the number of properties that are sold in a census tract. The concern is that drilling activity may affect the likelihood of a transaction, so that our sample of observed sales will be selected based upon the drilling exposure treatment. Using aggregated CoreLogic data, we regress the log of the annual number of transactions in each census tract on exposure to shale gas development within 20km of the tract centroid, including year and census tract fixed effects. We find that the effect of cumulative well pads is small and statistically insignificant for the number of properties sold (Table F.1). This is also true if we only focus on census tracts that are majority-piped-water areas or census tracts that are majority-groundwater areas. We therefore do not worry about sample selection in our housing transactions data induced by the well exposure treatment.

Table F.1: Log Number of Sales on Drilling Activity

	All Census Tracts	>50% PWSA Census Tracts	≥50% Groundwater Census Tracts
	(1)	(2)	(3)
	ln(# Sales)	ln(# Sales)	ln(# Sales)
Pads in 20km	3.77e-04	2.81e-04	5.87e-04
	(3.32e-04)	(3.87e-04)	(9.11e-04)
County-Year Effects	Yes	Yes	Yes
Census Tract Effects	Yes	Yes	Yes
n	19,283	16,353	2,930

Notes: Dependent variable is the log annual number of properties sold in a census tract, calculated using the property sales data. Each column represents a separate regression, differing based on the sample used: all census tracts in the data, census tracts that are mostly piped-water, and census tracts that are mostly groundwater. Regressor is the cumulative count of well pads drilled within 20km of the centroid of the census tract in the year of observation. Standard errors are clustered by census tract.

G Effects on Likelihood of New Construction

In this section, we perform two tests to investigate whether new construction associated with shale gas development may be driving down the size of the positive vicinity effect we find during the period around drilling. In particular, a strong increase in new housing supply may result in a failure to find any increase in prices in spite of a positive vicinity effect. Using CoreLogic data, we check first to see if the likelihood of a transaction for a newly constructed property is a function of exposure to cumulative well pads within 20km at the time of sale.[7] In particular, we run a regression at the property level, where the dependent variable is equal to one if the sale refers to a newly constructed house, and zero otherwise; the regression includes the count of well pads within 20km from the census tract, the count interacted with groundwater, census tract fixed effects and year fixed effects. Results are reported in Column (1) in Table G.1—we find that cumulative well pads are weakly negatively

[7] Whereas we had dropped new construction homes from our previous analyses, we reintroduce them to the dataset here. If we were to include newly constructed homes in our previous analyses, our findings would not change.

correlated with the likelihood of a transaction being a new construction.

Table G.1: New Construction on Drilling Activity

	Using All Property Sale Data
	Indicator (New=1)
Pads in 20km	-4.0e-04*
	(2.2e-04)
(Pads in 20km)*GW	2.5e-04
	(1.5e-04)
Census Tract Effects	Yes
County-Year Effects	Yes
Quarter Effects	Yes
n	634,820

Notes: The sample includes all properties sold in the property sales data; dependent variable equals 1 if the property was a new building, zero otherwise. Regressor is the count of wellbores (or well pads) that have been drilled within 20km of the property at the time of sale.

H Heterogeneity across Geography

The largest population center above the Marcellus shale in Pennsylvania is found in the Pittsburgh metropolitan area in the Southwest. Here we investigate whether the results are being driven by the Southwest, given that most of the properties in our sample are from this area.[8] In the first panel of Table H.1 we show that indeed the results from the Southwest subsample are very similar to the findings in our main table (Table 2). However, when restricting the sample to counties in the Northeast we also find somewhat similar results; properties that depend on private groundwater wells are negatively affected when in close proximity to shale gas development. We do not see a positive impact on PWSA properties when focusing on the Northeast and there are a couple of potential reasons for this. In the Northeast, 60 percent of the wells within 1.5km of properties sold in 2012 were not producing any gas, whereas in the Southwest, only 39 percent of the wells were not producing. Pipeline infrastructure is more developed in the Southwest making marginal wells more profitable, but perhaps more important for production is that in the western part of the Southwest, natural gas production is “wet,” meaning that alongside the methane production are natural gas liquids (ethane, butane, propane, and pentane). In the Northeast, the natural gas is “dry,” containing primarily methane. Over this time period, natural gas liquids have obtained a higher price than methane, making wells in the Southwest more profitable than in the Northeast. If we divide the sample

[8]Southwest counties included are Allegheny, Armstrong, Beaver, Butler, Fayette, Greene, Indiana, Lawrence, Somerset, Washington, and Westmoreland.

instead into counties with and without wet gas, the distinction between the estimates is even larger (Table H.2).[9] Areas with natural gas liquids have statistically significant increases in value when in close proximity to shale gas, even at a 1km distance. Properties in the PWSA boundary sample that are groundwater-dependent, see a smaller increase than PWSA properties. Properties in areas without natural gas liquids, do not see an increase in value, and those dependent on groundwater see a large decrease.

[9] We used a map with an approximation of the line dividing wet and dry Pennsylvania to designate counties as either with wet gas or without. See http://www.marcellus.psu.edu/images/Wet-Dry_Line_with_Depth.gif.

Table H.1: Log Sale Price on Well Pads by Geographic Subsamples

	$K \leq 1$km		$K \leq 1.5$km		$K \leq 2$km	
	Full	Boundary	Full	Boundary	Full	Boundary
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Properties in the Southwest						
Pads in Kkm	.027	.026	.029**	.035*	.016**	.018*
	(.025)	(.035)	(.014)	(.02)	(6.8e-03)	(.011)
(Pads in Kkm)*GW	-.043	-.162**	-.024	-.092**	9.0e-03	3.3e-03
	(.054)	(.075)	(.025)	(.038)	(.018)	(.053)
Pads in 20km	-7.9e-04**	-7.5e-04	-8.4e-04**	-8.9e-04	-8.5e-04***	-9.1e-04
	(3.2e-04)	(6.4e-04)	(3.3e-04)	(6.6e-04)	(3.3e-04)	(6.8e-04)
(Pads in 20km)*GW	6.7e-05	2.0e-03**	7.2e-05	2.0e-03**	-6.7e-05	1.6e-03*
	(5.8e-04)	(9.3e-04)	(6.0e-04)	(9.1e-04)	(6.1e-04)	(9.0e-04)
Property Effects	Yes	Yes	Yes	Yes	Yes	Yes
County-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes	Yes
n	199,344	52,986	199,344	52,986	199,344	52,986
p-value ($\alpha_3 + \alpha_4 = 0$)	0.766	0.075	0.814	0.174	0.119	0.692
Avg. Pads in Kkm	0.003	0.007	0.009	0.018	0.019	0.038
Avg. Pads in 20km	5.051	5.634	5.051	5.634	5.051	5.634
Panel B: Properties in Northeast						
Pads in Kkm	-5.9e-03	-.013	-.018	-9.2e-03	-3.5e-03	3.9e-03
	(.112)	(.115)	(.053)	(.064)	(.038)	(.039)
(Pads in Kkm)*GW	-.059	-.225	-.048	-.464**	-.08*	-.149*
	(.12)	(.194)	(.066)	(.233)	(.043)	(.088)
Pads in 20km	-1.2e-03*	-1.0e-03	-1.2e-03*	-1.0e-03	-1.2e-03*	-1.1e-03
	(6.3e-04)	(6.8e-04)	(6.5e-04)	(7.0e-04)	(6.7e-04)	(7.4e-04)
(Pads in 20km)*GW	1.7e-03**	2.0e-03*	1.9e-03**	2.3e-03**	2.3e-03***	2.5e-03**
	(7.2e-04)	(1.1e-03)	(7.5e-04)	(1.1e-03)	(8.4e-04)	(1.2e-03)
Property Effects	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes	Yes
n	28,068	11,762	28,068	11,762	28,068	11,762
p-value ($\alpha_3 + \alpha_4 = 0$)	0.287	0.225	0.128	0.047	0.002	0.043
Avg. Pads in Kkm	0.002	0.001	0.006	0.003	0.013	0.008
Avg. Pads in 20km	2.729	3.286	2.729	3.286	2.729	3.286

Notes: Each column in each panel represents a separate regression. Dependent variable in all regressions is the log sale price. Independent variables are the counts of wells at different distances from the property, drilled before the sale, as well as interactions with an indicator for whether the property is dependent on groundwater (GW). First panel only includes properties in the Southwest and the second panel only includes properties in the Northeast. The boundary sample restricts the full sample to include only properties in a narrow band around the border of the public water service areas. Robust standard errors are clustered by census tract. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table H.2: Log Sale Price on Well Pads by Geographic Subsamples: Abundance of Natural Gas Liquids

	$K \leq 1$km		$K \leq 1.5$km		$K \leq 2$km	
	Full (1)	Boundary (2)	Full (3)	Boundary (4)	Full (5)	Boundary (6)
Panel A: Properties in Wet-Gas Counties						
Pads in Kkm	.063**	.071**	.052***	.057**	.024***	.026**
	(.032)	(.03)	(.018)	(.022)	(9.1e-03)	(.012)
(Pads in Kkm)*GW	-.067	-.201***	-.027	-.11***	4.3e-03	-3.5e-03
	(.053)	(.075)	(.026)	(.037)	(.019)	(.054)
Pads in 20km	-1.3e-03***	-2.1e-03*	-1.3e-03***	-2.2e-03**	-1.3e-03***	-2.2e-03**
	(4.2e-04)	(1.1e-03)	(4.3e-04)	(1.1e-03)	(4.3e-04)	(1.1e-03)
(Pads in 20km)*GW	-1.0e-04	2.1e-03**	-1.4e-04	2.1e-03**	-2.5e-04	1.6e-03
	(6.3e-04)	(1.0e-03)	(6.6e-04)	(1.0e-03)	(6.7e-04)	(1.0e-03)
Property Effects	Yes	Yes	Yes	Yes	Yes	Yes
County-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes	Yes
n	165,421	42,126	165,421	42,126	165,421	42,126
p-value ($\alpha_3 + \alpha_4 = 0$)	0.937	0.102	0.216	0.225	0.076	0.688
Avg. Pads in Kkm	0.003	0.008	0.007	0.017	0.015	0.035
Avg. Pads in 20km	4.160	4.569	4.160	4.569	4.160	4.569
Panel B: Properties in Dry-Gas Counties						
Pads in Kkm	-.026	-.121	-6.1e-03	-.024	1.7e-03	-2.7e-03
	(.035)	(.073)	(.01)	(.026)	(5.8e-03)	(.016)
(Pads in Kkm)*GW	-.033	-.151	-.074*	-.421*	-.078***	-.121*
	(.079)	(.197)	(.041)	(.214)	(.026)	(.07)
Pads in 20km	-6.0e-05	1.6e-04	-6.1e-05	1.2e-04	-1.0e-04	3.0e-05
	(4.1e-04)	(5.0e-04)	(4.1e-04)	(5.1e-04)	(4.1e-04)	(5.6e-04)
(Pads in 20km)*GW	1.3e-03**	1.7e-03*	1.5e-03**	2.0e-03**	1.8e-03***	2.1e-03**
	(5.7e-04)	(9.5e-04)	(5.9e-04)	(9.8e-04)	(6.5e-04)	(1.0e-03)
Property Effects	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract-Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Quarter Effects	Yes	Yes	Yes	Yes	Yes	Yes
n	64,525	24,201	64,525	24,201	64,525	24,201
p-value ($\alpha_3 + \alpha_4 = 0$)	0.414	0.141	0.049	0.037	0.003	0.063
Avg. Pads in Kkm	0.004	0.003	0.012	0.011	0.026	0.025
Avg. Pads in 20km	6.174	6.047	6.174	6.047	6.174	6.047

Notes: Each column in each panel represents a separate regression. Dependent variable in all regressions is the log sale price. Independent variables are the counts of wells at different distances from the property, drilled before the sale, as well as interactions with an indicator for whether the property is dependent on groundwater (GW). First panel only includes properties that are in counties that have natural gas liquids and the second panel only includes properties in counties without any natural gas liquids. The boundary sample restricts the full sample to include only properties in a narrow band around the border of the public water service areas. Robust standard errors are clustered by census tract. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

References

Andersson, H., L. Jonsson, and M. Ögren (2010). Property prices and exposure to multiple noise sources: Hedonic regression with road and railway noise. *Environmental and Resource Economics 45*(1), 73–89.

Bajari, P. and C. L. Benkard (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy 113*(6), 1239–1276.

Bajari, P., J. Cooley, K. il Kim, and C. Timmins (2010). A theory-based approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution. *American Economic Review 102*(5), 1898–1926.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *The Journal of Political Economy 95*(1), 81–88.

Bishop, K. and C. Timmins (2012). Hedonic prices and implicit markets: Estimating marginal willingness to pay for differentiated products without instrumental variables, Duke University Working Paper.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics 114*(2), 577–599.

Brown, J. and S. Rosen (1982). On the estimation of structural hedonic price models. *Econometrica 50*(3), 765–768.

Bui, L. T. and C. J. Mayer (2003). Regulation and capitalization of environmental amenities: Evidence from the Toxic Release Inventory in Massachusetts. *Review of Economics and Statistics 85*(3), 693–708.

Chay, K. Y. and M. Greenstone (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy 113*(2), 376–424.

Davis, L. W. (2004). The effect of health risk on housing values: Evidence from a cancer cluster. *The American Economic Review 94*(5), 1693–1704.

Davis, L. W. (2011). The effect of power plants on local housing values and rents. *Review of Economics and Statistics 93*(4), 1391–1402.

Ekeland, I., J. Heckman, and L. Nesheim (2004). Identification and estimation of hedonic models. *Journal of Political Economy S1*, S60–S109.

Epple, D. (1987). Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Journal of Political Economy 95*(1), 59–80.

Gamper-Rabindran, S. and C. Timmins (2013). Does cleanup of hazardous waste sites raise housing values? evidence of spatially localized benefits. *Journal of Environmental Economics and Management 65*(3), 345–360.

Greenberg, M. and J. Hughes (1992). The impact of hazardous waste superfund sites on the value of houses sold in New Jersey. *The Annals of Regional Science 26*(2), 147–153.

Greenstone, M. and J. Gallagher (2008). Does hazardous waste matter? Evidence from the housing market and the superfund program. *The Quarterly Journal of Economics 123*(3), 951–1003.

Greenstone, M. and R. Hanna (2014). Environmental regulations, air and water pollution, and infant mortality in india. *American Economic Review 104*(10), 3038–3072.

Haninger, K., L. Ma, and C. Timmins (2012). Estimating the impacts of brownfield remediation on housing property values, Duke University Working Paper.

Harrison Jr, D. and D. L. Rubinfeld (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management 5*(1), 81–102.

Heckman, J., R. Matzkin, and L. Nesheim (2010). Nonparametric Identification and Estimation of Nonadditive Hedonic Models. *Econometrica 78*(5), 1561–1591.

Kiel, K. A. and M. Williams (2007). The impact of Superfund sites on local property values: Are all sites the same? *Journal of Urban Economics 61*(1), 170–192.

Kuminoff, N. V. and J. Pope (2014). Do "Capitalization effects" for public goods reveal the public's willingness to pay. *International Economic Review 55*(4), 1227–1250.

Leggett, C. G. and N. E. Bockstael (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management 39*(2), 121–144.

Linden, L. and J. E. Rockoff (2008). Estimates of the impact of crime risk on property values from Megan's Laws. *The American Economic Review 98*(3), 1103–1127.

Linn, J. (2013). The effect of voluntary brownfields programs on nearby property values: Evidence from Illinois. *Journal of Urban Economics Forthcoming*.

Mendelsohn, R. (1985). Identifying Structural Equations with Single Market Data. *The Review of Economics and Statistics 67*(3), 525–529.

Palmquist, R. B., F. M. Roka, and T. Vukina (1997). Hog operations, environmental effects, and residential property values. *Land Economics*, 114–124.

Poor, J. P., K. L. Pessagno, and R. W. Paul (2007). Exploring the hedonic value of ambient water quality: A local watershed-based study. *Ecological Economics 60*(4), 797–806.

Pope, J. C. (2008a). Buyer information and the hedonic: the impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics 63*(2), 498–516.

Pope, J. C. (2008b). Fear of crime and housing prices: Household reactions to sex offender registries. *Journal of Urban Economics 64*(3), 601–614.

Ridker, R. G. and J. A. Henning (1967). The determinants of residential property values with special reference to air pollution. *The Review of Economics and Statistics 49*(2), 246–257.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy 82*(1), 34–55.

Walsh, P. J., J. W. Milon, and D. O. Scrogin (2011). The spatial extent of water quality benefits in urban housing markets. *Land Economics 87*(4), 628–644.

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