Reading Comments - Environmental Economics II

Jeanne Sorin

January 16, 2021

Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, Rosen 1974

Summary

This theoretical paper is the foundational paper on hedonic prices analysis, which is very popular in many economic fields, including in urban economics when the analysis is run on real estate prices. Hedonic prices are "*implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them*". Differentiated products are goods that can be fully described by their vector of characteristics. The model at the basis of hedonic prices analysis is a model of competitive equilibrium, in a plane of as many dimensions as good characteristics $z = (z_1, z_2, \ldots, z_n)$, where buyers and sellers meet. Prices $p(z)$ are functions off all zi characteristics.

Consumers get utility from each zi characteristic, their utility function is strictly concave. The first order conditions of the utility maximization problem give $pi/px = (dU/dz)/(dU/dx)$. It is interesting to approach the problem by defining a bid function $\theta(z_1, \ldots, z_n, u, y)$ where $u = U(y - \theta, z_1, \ldots, z_n)$, which is the expenditure a consumer is willing to pay for an alternative z, given y and u and defines a series of indifference surfaces (equivalent to indifference curves in multiple dimensions). One can show that only if the price function p(z) is convex, higher income consumers consume more of each good. However, in general, no such claim can be made. Besides, Rosen emphasizes that in general, once must impose a lot of structure for the problem to have a solution, and even more for it to have a closed-form solution. Very interestingly, the model predicts market segmentation in terms of tastes (value function), as in Tiebout's (1956) implicit neighborhood market analysis. This suggests the potential role of hedonic analysis for spatial models.

Symmetrically, on the supply side, different firms have different cost functions, and therefore different offer functions $\phi(z_1, \ldots, z_n, \pi, \beta)$, which is the unit price firms are willing to accept to produce an alternative z for a given profit π and shift parameter β.

The observed $p(z)$ is the market clearing price function, enveloping a family of value functions theta and offer functions ϕ. The tangency of the value and offer functions lead to equilibrium, i.e. $Q_d(z) = Q_s(d)$ at $p(z)$. Rosen discusses several types of market equilibrium: short run equilibrium where z is fixed and only Q and p can change ; short run equilibrium where p is fixed and only Q and z can change; long-run equilibrium where Q, p and z can

all change. In the latter case, $p(z)$ is only a function of supply, because it is determined by minimum average cost of z.

In order to estimate the model, as the observed price is an equilibrium object such that the marginal demand price for zi is equal to the marginal supply price for zi, for all zi, and the price for each zi is not observed, one must run a two-step procedure. In a first step, p(z) is estimated using the usual hedonic method, and then compute the implicit marginal price $dp(z)/dzi$. In a second step, use the estimated implicit marginal price as the endogenous variable for each system of simultaneous equations (for each zi)

Finally, Rosen discusses how this approach can be used to investigate the welfare consequences of quality-standard regulations.

Criticisms and Questions

I do not understand the underlying reasoning behind "*also, as a general methodological point, it is demonstrated that conceptualizing the problem of product differentiation in terms of a few underlying characteristics instead of a large number of closely related generic goods leads to an analysis having much in common with the economics of spatial equilibrium and the theory of equalizing differences*".

Besides, the role of arbitrage activities (p37) in forcing p(z) to be linear, is unclear to me. More precisely, I don't understand why a linear price is problematic, as linear functions are both convex and concave. More generally, I would have appreciated more extensive discussion about the role of linearity vs nonlinearity. For example: how relevant empirically are linear vs nonlinear price functions? In which domain?

Finally, having seen Rosen (1974) cited many times in urban / housing economic papers, I was expecting this paper to be more explicitly about hedonic analysis in the real estate market context, and found it quite striking how much more general it is.

Does Hazardous Waste Matter? Evidence From The Housing Market And The Superfund Program, Greenstone & Gallagher 2008

Summary

This paper investigates whether individuals care about hazardous waste, and more precisely the extent to which the housing market appreciates in response to Superfund-sponsored cleanups of hazardous waste sites. This program, launched in 1985 by the Environmental Protection Agency (EPA), placed dangerous waste sites on a National Priorities List (NPL) so that they could be cleaned, starting in 2005. This program was controversial because of his cost and delays, and this paper moreover suggests that its welfare consequences were not substantial. The authors use the evolution of housing market outcomes as a proxy for the latter.

The basis to estimate the value of local changes in environmental amenities is the hedonic model from Freeman (1974) and Rosen (1974), where a differentiated good (a house) is defined by a list of characteristics. The price of each of them equals the partial derivative of the good's price with respect to the relevant characteristic in equilibrium. This model predicts that if inhabitants care about the cleanups, we should observe an increase in both supply and demand of housing, leading to an increase in both prices and quantity, as well as sorting of individuals caring relatively more about such environmental amenity into cleaned sites.

However, successfully recovering individuals' valuations of environmental amenities (e.g. clean air, clean ground etc) through a revealed preferences approach, like housing prices, is subject to many misspecification threats. In the case of the superfund program, a standard hedonic approach would suffer from a selection bias, as sites listed in the NPL are arguably not comparable to other sites. The balance test on all control vs treated census tracts fails, suggesting that a selection bias exists, which would bias the OLS estimates. In order to address this selection bias issue, the authors exploit the (known) selection criteria for sites to be added to the NPL, as well as the sequential selection of sites to be examined. In a first step, 14,697 sites were referred to the EPA to be examined. In a second step, 690 out of 14,697 sites were identified as very dangerous. Because of cost concerns, the EPA was then able to add only 400 of these 690 sites to the NPL. In order to make this selection, sites got assigned a Hazardous Ranking System (HRS) score from 0 to 100 depending on their toxicity and proximity to humans. Sites with a score above 28.5 made it to the NPL. The existence of first a shortlist, and only then the calculation of the HRS and the resulting cutoff is the main lever of this paper, as it arguably provides a valid research design where the 690 sites are split at the cutoff. The 28.5 cutoff being only an imperfect predictor of treatment, it is used both as an instrument in a standard IV analysis (to account for potential endogenous rescoring), and in a quasi-experimental RD design. The balance test on census tracts surrounding sites with 1982 HRS scores is much more robust. These two approaches are compared to a conventional hedonic approach looking at the housing market of census tracts with sites in the NPL vs the housing market of other (surrounding) census tracts. The conventional approach relies on the assumption that unobserved determinants of housing prices are uncorrelated with NPL status.

These analyses are run on a dataset containing (1) detailed information about 1,398 haz-

ardous sites, (2) housing, demographic and economic panel from Geolytics' Neighborhood Change Database at the Census tract level. Levels of analysis are (1) census tracts, (2) census tracts sharing a border with tracts containing a hazardous site and (3) land area within circles of different radii around these sites.

According to the conventional approach (Table III), the cleanups led to a statistically significant increase in housing prices from 4.0 to 19.1%. This holds both on the full census tract sample and on a sample restricted to NPL sites' and neighboring census tracts. However, both the IV and the quasi-experimental RD designs (Table IV) lead to a different conclusion, as changes in housing prices after the cleanups are small and not statistically significant at conventional levels. The same conclusion is obtained from looking at rental rates rather than housing prices. With regards to both housing supply and sorting, the null hypothesis of a zero impact cannot be rejected.

The authors propose three potential explanations for these results. First, a low WTP for clean sites in by populations choosing to live near them both before and after the cleanups. Second, it could be the case that the population doesn't believe that cleanups significantly improve the environmental amenity. Third, control units could have actually also been somewhat treated through remediation activities. The first two hypotheses are preferred by the authors.

These findings are important from a policy perspective, as the cost of the cleanups is estimated at $43 million which, if the housing market captures the full benefit of the program, can be considered as a "waste".

Criticisms and Questions

I really like how the paper introduces the empirical strategies addressing the issues brought by the conventional hedonic analysis. Estimating the latter, in addition to the main analysis, gives a lot of force to the alternative strategies, as the reader can see by himself how biased can the conventional approach be.

However, I am not so convinced by the discussion on whether the housing market really encompasses all impact of the cleanup. From my understanding of the paper, this is equivalent to assuming that the housing supply curve is perfectly inelastic. Do we have any way to check this, maybe using Zillow ZTRAX data? Without more evidence, I find it hard to make the claim that there was no welfare gain at all from the fact that the housing market didn't react. If, however, the cleanups have no effect on actual environmental amenities, as suggested on page 995, such claim is more convincing. Proving this no effect is beyond the scope of this paper.

Finally, I am not sure to fully understand what exactly is the "control function" solution brought by RD compared to IV.

The estimation of Demand Parameters in Hedonic Price Models, Bartik 1984

The main point of this paper is that the problem of estimating hedonic demand parameters is caused not by demand-supply interac- tion but by the endogeneity of both marginal prices and quantities when households face a nonlinear budget constraint. Appropriate instrumental variables for this problem should exogenously shift the hedonic price function; previously proposed instruments do not do so. The practical problem for empirical hedonic research is finding instruments whose exogeneity can be defended with some plau sibility.

A Unified Framework for Measuring Preferences for Schools and Neighborhoods, Bayer, Berreia, McMillan 2007 JPE

Summary

This paper estimates household's Marginal Willingness To Pay (MWTP) for school quality and neighborhood attributes. More generally, this paper proposes a framework using Boundary Discount Discontinuity to consistently estimate household preferences for some spatial amenity in the presence of sorting.

The main lever of this paper is its identification strategy and emphasis on sorting. The authors combine insights from hedonic price analysis with a discrete location choice model and exploit detailed data on school attendance boundaries. The latter creates an exogenous discontinuity in school quality.

The story could stop here if one was to compare houses immediately on each side of the boundary. However, in practice, he exercise is made harder by the necessity (for statistical power) to compare houses in the neighborhood (0.2, 0.3 miles) of the boundary, potentially adding compositional effects in the case of sorting. Following evidence of such sorting in terms of socio-demographic variables, the authors modify the canonical hedonic price approach to include a boundary dummy, capturing unobserved neighborhood characteristics[1].

The main findings from the hedonic analysis are the following. First, neighborhood sociodemographics have an impact on house prices, even after controlling for school quality and unobservables (through a boundary dummy). This suggests that controlling for the correlation between observable and unobservable neighborhood quality is key for not overestimating the MWTP for neighborhood socioeconomic characteristics. Besides, while house prices and neighborhood race are negatively correlated, the coefficient drops to insignificant levels once one controls for the correlation between neighborhood race and unobserved neighborhood quality (boundary dummy) (cols 1,3 vs 2,4 of Table 3). Furthermore, as expected, the MWTP for school quality is positive, but significantly decreases once one controls for unobserved neighborhood characteristics.

In response to evidence on sorting, the authors estimate a Random Utility Model (RUM) of residential location decision (discrete choice model), where households choose the house h giving them the maximum net indirect utilities. The model is estimated following a two-step procedure. In a first step, a MLE recovers estimates for the heterogeneous parameters and mean indirect utility. In a second step, heterogeneous parameters are decomposed into observables and unobservable components.

Criticisms and Questions

While I found the paper a little convoluted at first, a careful read has made me really enjoy it. One of my favorite aspect of the paper (about which I'm still slightly confused) is the discussion on the interpretation of the estimated coefficients of the hedonic price regression in the presence of sorting. Indeed, the authors emphasize that if a good is abundant, then the mean preferences recovered from hedonic price regression is likely to be very close to the MWTP. However, if a good (demand curve   MWTP curve is downward sloping) is

---

[1]The boundary dummy is in addition to a variable for school quality. As a reminder, canonical RDD does **not** include dummies for boundary, but rather simply study a buffer around the boundary.

scarse, the equilibrium price, recovered from the hedonic regression, will likely overestimate the mean MWTP of households for the amenity (recovered through the sorting model).

Besides, Table 6 summarizing the estimation procedure and key identifying assumptions is really useful.

I also found the result that race not being priced into house prices does not involve that there is no sorting on race because "*a sorting equilibrium can be achieved without a race being capitalized into houseing prices*" extremely interesting. This finding raises a lot of questions regarding how not to misinterpret zero hedonic price coefficients as lack of MWTP for an "amenity".

However, I have found other parts confusing. In Section II.C Data - School Attendance Zone Boundaries, the authors mention that for each block, they "*located the "twin" census block on the other side of that boundary*". However, I don't understand well the point of this matching. Are we including FE for each pair? Are we just making sure that there are the same number of blocks for each distance on each side of the boundary?

Besides, I believe the paper would have benefited from a more extensive discussion of the potential dynamics of price adjustment in response to changes in school quality. My understanding is that the authors rely on the 1990 Decennial Census, housing transactions between 1992 and 1996 and school test scores. However, I couldn't find (in the main text at least) any discussion on whether test scores changed over time, whether test scores in 1990 (?) were the ones taken into considerations by buyers in 1990 onwards, or whether then test score trends also matters.

Finally, I found section V.B Sorting Model - Estimation, fairly confusing.