Lecture 2: NBERMetrics.

Ariel Pakes

July 17, 2012

Topics.

- Sources of identification from market Level Data.
- Confronting the precision problem.
- Adding the Pricing Equation.
- The pricing equation and instruments.
- The pricing equation and hedonic models.

Sources of Identification on Product Level Data.

Lessons from Estimation: Market Level Data.

Often we find that there is not enough information in product level demand data to estimate the entire distribution of preferences with sufficient precision. The extent of this problem depends on the type of data. If there is a national market we typically have data over time periods, and when there are local markets the data is most often cross-sectional (though sometimes there is a panel dimension). Thus the researcher is trying to estimate the whole distribution of preferences from one set of aggregate choice probabilities per market or per period depending on the type of data. Other than functional form, the variance that is used to estimate the parameters of the model becomes

- differences in choice sets across markets or time periods (hopefully allowing you to sweep out preferences for given characteristics), and
- differences in observable purchaser characteristics across markets or over time (usually due to demographics) over a fixed choice set (hopefully allowing you to sweep out the interaction between demographic characteristics and choice sets).

It is easy to see how precision problems can arise. In the national market case (as in the original BLP, 1995, auto example), the distribution of purchaser characteristics rarely changes much over the time frame available for the analysis. As a result, one is relying heavily on variance in the choice set, which depending on the important characteristics of the product, is often not large (though one can think of products where it is, at least in certain characteristics, e.g. computers).

In the cross sectional case there is often not much variance in the choice set, though this does depend on the type of market being studied. The choice sets for markets for services are typically local and do often differ quite a bit with market characteristics. When there is variance in choice sets across markets one must consider whether the source of that variance causes a selection problem in estimation. In the terminology of the last lecture we must consider whether $\xi_{k,j}$, where k is a product and j is a market, is not a draw from a random distribution but is a draw from a distribution which depends on market characteristics. If so, there should be an attempt to control for it; i.e. model the distribution of ξ as a function of the relevant market characterisitics.

As was shown above there is often quite a bit of variance in demographic characteristics across markets, but there are often sources of variance that are important for the particular choice which we do not have data on. We do typically have data on income, family size, ect., but we will not have data on other features, like household activities or holdings of related products, which are often important to the choice. For example, in auto choice it probably matters what type of second car the family has, and whether the household has a child on a sports team where some fraction of the team must be driven to games periodically. These unobserved attributes will go into the variance of the random coefficients, and if their distribution differs by market we will have to parameterize them and add to the parameters that need to be estimated.

Solutions to the Precision Problem.

If there is a precision problem, to get around it we have to add information. Fundamentally there are two ways of doing this

- add data, and
- add assumptions which allow you to bring to bear more information from existing data.

Add Data. There are at least two forms of this. One is to add markets, and this just makes the discussion we have already had on sources of identification more relevant. The other is to add micro data. This will be the topic of my next lecture as there are differnt types of micro data and additional estimation issues arise when dealing with each type.

What makes the incorporation of micro level data relatively easy for this class of models is that the model we just developed for market level data is a micro model (we just aggregated it up), so *the model used for the micro data will be perfectly consistent with the models just developed for market level data.* Though micro data sets, data sets that match individuals to the choices those individuals make, is increasingly available, it is not as available as market level data. Indeed typically if micro data is available, so is aggregate data, and the models and estimation techniques we develop and use for micro models will be models that incorporate informaton on both micro and market level data.

For now it suffices to say that the additional micro data sometimes available can come in different forms, typical among them:

- There are publically available purchaser samples that are too small to use in estimating a rich micro model but large enough to give fairly precise estimates of covariances between household and product characteristics which an adequate aggregate model should replicate. A good example of this is the Consumer Expenditure Survey (or the CEX), and we will return to this when we look at Petrin's work. See also Aviv's discussion of Goldberg's paper.

- There are privately generated surveys that tell you something about the average characteristics of consumers who buy different products. These are often generated by for-profit marketing firms. As a result, it can be costly to access current data, but these companies will often cut deals on old data if they are not currently selling it.

- True micro data which matches households to the products they chose. This is still fairly rare in Industrial Organization but is becoming more widely available in public finance applications (e.g. schooling choices). This type of data is particularly useful when it contains not only the product purchased but second and other ordered choices. We come back to it when we discuss the CAMIP data and MicroBLP, and when we discuss the Homescan data which is used Katz's study (which I will turn to in the next lecture), and much of the other work on purchases at food outlets. It is a panel; which does not quite give us ordered choice, but has some similar features (see the micro lecture).

Adding assumptions. There are a number of ways of doing this, but probably the most prevalent is to add a pricing assumption. This was done in BLP and in much of the literature preceeding them. The assumption is typically that the equilibrium in the product market is "Nash in prices", sometimes called Bertrand.

To implement it we assume a cost function and add the equation for the Nash pricing equilibrium to the demand equation to form a system of equations which is estimated jointly (this is reminiscent of old style demand and supply analysis; with the pricing equation replacing the supply function). Since there is a pricing equation for each product, there is a sense in which the pricing assumption has doubled the amount of observations on the left-hand side variable (though, as we shall see we expect the errors in the two equations to be correlated). Of course typically the use of the cost function does require estimates of more parameters, but just as we did for the demand function, we typically reduce the cost function to be a function of characteristics (possibly interacted with factor prices) and an unobservable "productivity" parameter. So the number of paramaters added is a lot less than the number of products, and we get a pricing equation for each product.

The assumption of a Nash equilibrium in prices can be hard to swallow, as there are many markets where it would seem to be *a priori* unreasonable. Examples include all markets where either demand or costs are "dynamic" in the sense that a change in current quantity changes future demand or cost functions. We expect this to be true in durable, experience or network goods, as well as in goods where there is learning by doing or important costs of adjusting quantities supplied. What is surprising is how well the Nash pricing assumptions fit in characteristic based models, even in markets where one or more of these issues is relevant.

There is, however, an empirical sense in which this should have been expected. Its been known for some time that if we fit prices to product characteristics we get quite good R^2's, as it was this fact that generated the hedonic's literature (see the discussion and illustration below). Since marginal costs are on the right had side of the Nash pricing assumption, and marginal costs are usually allowed to be a function of characteristics (perhaps in conjunction with factor prices) this already insures that the pricing function will have a reasonable fit.

The other term on the right-hand side of the pricing equation is the markup. The markup from the Nash pricing assumption has properties that are so intuitive that they are likely to also be properties of more complex pricing models. In particular, the Nash pricing model implies:

- Products in a heavily populated part of the characteristic space will have high price elasticities; if their prices are increased, there are lots of products which are close in characterisitc space to switch to and consumer's utility is only a function of the characteristics of the product. High elasticites means low markups (see below).

- Products with high prices are sold to less price sensitive (or high income) consumers and hence will have lower elasticities and higher markups.

- If a firm is multi-product the markup it charges on any one of its products is likely to be higher than it would have been had the firm been a single product firm (when a multi-product firm increases its price, some of the consumers that leave the good go to the other good the firm owns and the firm does not lose all the markup from the first good those consumers generated).

The fact that the markup term is a function of price elasticities is also what endows that equation with highly relevant information on the structure of the demand system.

“Bertrand” Equilibrium in a Differentiated Product Market, (the Nash in Prices Solution with Multiproduct Firms) .

Since most of the markets we study involve multiproduct firms, we investigate price-setting equilibria in differentiated products market with multiproduct firms. The standard assumption is that the price setter is maximizing total firm profits. Though I comment below on alternatives, we begin with this assumption. Keep in mind when we use this assumption in estimation, we are actually adding two assumptions to our demand model

- an assumption on the nature of equilibrium and

- an assumption on the form of the cost function.

The Cost Function. A frequently used specification for the marginal cost is

$$ln[mc_j] = \sum_r w_{r,j}\gamma_k + \gamma_q q_j + \omega_j,$$

where the $\{w_{r,j}\}$ typically include product characteristics (the $\{x_{k,j}\}$ and factor prices (often interacted with product characteristics). The ω_j are unobserved determinants of costs, and q_j picks up the possible impacts of non-constant returns to scale.

Keep in mind that the ξ_j are not included in this equation, so if it costs to produce these "unobserved" characteristics we would expect ξ_j to be positively correlated with ω_j. Of course ω_j also includes the impact of productivity differences across firms. They are typically assumed to be i.i.d. draws from a population distribution which is indpendent of the $\{w_{r,j}\}$ and has a mean of 0. Notice that then the distribution of marginal cost depends on only $R+1$ parameters, and we have J observations on price (hopefully $J >> R$), so we are adding degrees of freedom). Also keep in mind that q_j is endogenous, in the sense that it is a function of both the ξ_j and the ω_j; both unobserved cost and unobserved product components impact demand.

This function is often also assumed (at least formally incorrectly) to pick up other aspects of the environment that we can not deal with in a static model. For example if some of the inputs are imported, we should be able to evaluate changes in costs of those products due to exchange rate changes directly. Typically though these exchange rate changes are not passed through to consumers in full (this is the exchange rate pass through literature in trade). Instead of specifying the formal price setting reason for not passing through the exchange rate fluctuation in full, much of the literature just puts the exchange rate in the cost function, estimating a "pass through rate".

The Pricing Equation. For simplicity I am going to assume constant costs (or $\gamma_q = 0$). Since the equilibrium is Nash in prices for a profit maximizing firm, if J_f denotes the set of products the firm markets, their price choices must

$$max_p \pi_f(\cdot) = \sum_{j \in J_f} (p_j - mc_j) M s_j(\cdot).$$

Then the f.o.c. for each of the $\#J_f$ prices of products owned by the firm is

$$s_j(\cdot) + \sum_{r \in J_f} (p_r - mc_r) M \frac{\partial s_r(\cdot)}{\partial p_j} = 0,$$

and for these to be equilibrium prices second order conditons must also hold.

Notice that for a single product firm the f.o.c. reduces to

$$p_j = mc_j + \frac{s_j(\cdot)}{-\partial s_j(\cdot)/\partial p_j}$$

which is the familiar formula for price; cost plus a markup which equals one over the (semi) elasticity of demand.

When there are two products owned by firm j say good 1 and 2, this becomes

$$p_{1,j} = mc_{1,j} + \frac{s_{1,j}(\cdot)}{-\partial s_{1,j}(\cdot)/\partial p_{1,j}} + \frac{\partial s_{2,j}(\cdot)/\partial p_{1,j}}{-\partial s_{1,j}(\cdot)/\partial p_{1,j}} [p_{2,j} - mc_{2,j}].$$

This illustrates the intuition noted earlier. Since we are dealing with a differentiated product market, goods are substitutes and the last term is positive. So if we held all other product prices fixed for the comparison, prices will be higher for multi-product firms. The extent to which they will be higher depends on the cross price elasticity of good 2 with the price of good 1, and the markup on good 2 (and there is a similar reasoning for good 2).

Of course, if prices are really equilibrium prices, once the price of either good 1 or good 2 changes, then the prices of the other goods in the market will also adjust (as they will no longer satisfy their first order conditions) and we could not get the full effect of going from a single to a double product firm without calculating a new equilibrium. As a further caveat to thinking about counterfactuals in this way, note that if prices are set in a simultaneous move game, then there may be more than one equilibrium price vector: more than one vector of prices that satisfies all the first (and second) order conditions. All we will rely on in estimation will be the first order condition, so the potential for multiplicity does not affect the properties of the estimators we introduce. However the calculation of counterfactual prices, the prices that would hold were we to change the institutional structure or break up a multiproduct firm, does depend on the equilibrium selection mechanism and without further information (like an equilibrium selection procedure) or functional form restrictions (say, that insure a unique equilibrium) can not be calculated.

Other Pricing Assumptions. This is not the only pricing assumption that is possible. One might want to see if a division of a firm that handled a subset of its products was pricing in a way that maximized the profits from that subset rather than from the products of the whole firm or if two firms were co-ordinating their pricing decisions and trying to maximize the sum of their profits. Each of these (and other) assumptins would deliver a different pricing equation, and the new equations could replace this equation in the estimation algorithm I am about to introduce. Similarly one could try and set up an algorithm which tested which pricing assumption was a better description of reality. For an example of this, see Nevo (Econometrica, 2001).

Details. Define the matrix Δ which is $J \times J$ where the (i, j) element is

- $\partial s_i(\cdot)/\partial p_j$ if both $i \in J_f$ and $j \in J_f$ for some firm f
- and the (i, j) element is zero if the two goods do not belong to the same firm.

Then we can write the vector of f.o.c.s in matrix notations as

$$s + (p - mc)\Delta = 0 \quad \Rightarrow \quad p - \Delta^{-1}s = mc.$$

Turning to estimation, now *in addition* to the demand side of J equations for ξ we have the pricing side equations for ω which can be written as

$$ln(p - \Delta^{-1}s) - w'\gamma = \omega(\theta).$$

Recall that once we isolated ξ the assumption that ξ was orthogonal to the instruments allowed us to form a moment which was mean zero at the true parameter value. Now we can do the same thing with ω. I.e. if z is an insturment then

$$Ez_j\omega_j(\theta) = 0 \;\; at \;\; \theta = \theta_0.$$

How much have we complicated the estimation procedure?

- Before, everytime we wanted to evaluate a θ we had to simulate demand and do the contraction mapping for that θ. Now after simulating demand we have to calculate the markups for that θ also.
- Now we have twice as many moment restrictions to minimize, and this will often increase computing time somewhat.

Some Final Notes on Estimating the Pricing Equation. When we estimate the pricing equation with the demand system we require more assumptions than when estimating the demand system alone; and some of them are problematic. On the other hand

- As noted the pricing equation often provides a lot of information, and whether or not it is exactly right, it is a model that fits quite well.
- Also, you will often (though not always) need a variant of pricing assumptions for policy analysis anyway, and so it is not clear why we should not impose it right off.

On the latter point there are at least three possiblities for subsequent use.

- There are some issues which don't require a pricing assumption at all; e.g. analyzing actual consumer surplus changes over time, as is required in the constructing of price indices; see below.

- There are some issues for which we might think that the impact of the pricing assumption is second order; e.g. the impact of a tax on gas, or of taking a good off the market. Both of these will change equilibrium prices but we might think that the effects of the equilibrium price change is second order.

- Issues when the whole purpose of the analysis is the equilibirum price change; like merger analysis, or the welfare impact of policies that change prices.

Pricing Equations and The Choice of Instruments.

If we are willing to assume a conditional moment restriction; i.e. that our unobservables are mean zero conditional on some pre-specified variables, this generates a lot of instruments (any function of the pre-specified variables will do). However, under mild regularity conditions, Chamberlain (*Journal of Econometrics*, 1987) has shown that there are optimal instruments, in the sense that the use of these instruments will minimize the variance of the estimator over any possible function of the pre-specified variables. The number of these instruments is exactly equal to the number of parameters being estimated. To the extent we can, then, we might attempt to use the "optimal" instruments as a way of improving precision.

There are two problems with this approach, but neither truly undermines it. The problems are

- to derive the optimal instrument formula we will have to make assumptions on the pricing equation (if product characteristics were endogenous we would also need an equation describing how those are set),
- the optimal instrument formula generates an instrument which is hard to compute.

The reason these two problems do not undermine the approach is that if we miss-specify the pricing equation, or if we use an easy to compute approximation to the true instrument, the estimator we derive will *still be consistent.* Moreover the loss of efficiency relative to using the optimal instrument will depend on the closeness of the approximation to the optimal instrument formula. With regards to the pricing formula, we have already noted that though we may not believe that the Nash in prices formula is exactly correct, we do know that it produces a very good approximation to prices.

To illustrate I will use the conditional moment restriction used in BLP (and Aviv will give you alternatives in the next lecture). Foreshadowing the notation I will use later, let β^o be the parameters from the projection of δ on product characteristics and β^u be the parameters on the random coefficients. If we are willing to assume $E[\xi(\beta^o, \beta^u)|x, w] = 0$ where here (x, w) represent the matrix of all firms characteristics and factor prices, then the optimal moments will be

$$\frac{\partial E[\xi(\beta^o, \beta^u; s)|x, w]}{\partial \beta} = \frac{E\partial[\delta(\beta^u; s) - x\beta^o_x - p\beta^o_p|x, w]}{\partial \beta}$$

where it is understood that

- δ is obtained from inverting the actual shares (s) when the random coefficients are set at β^u, and
- all derivatives are evaluated at the true value of the parameter vector.

The derivative w.r.t.

β^o_x is just x; $\quad \beta^o_p$ is $E[p|x, w]$; $\quad$ and β^u is $E[\partial\delta(\beta^u; s)/\partial\beta^u|x, w]$.

Since the instruments depend on the true value of the parameter vector, this becomes a two-step estimator. In the first step we use an inefficient set of instruments to obtain an initial consistent estimate of the parameter vector. In the second step we use the first stage estimate to calculate approximations to the last two derivatives, and then use those approximations as instruments in the estimation algorithm. This is an "adaptive" two step estimator in the sense that the estimation error in the first round does not contribute to the variance of the asymptotic distribution of the second round.

The last two derivatives could not be computed directly from the data. BLP(1999, *AER*) approximate them by calculating the Bertrand price equilibrium setting at $\xi_j = \omega_j = 0, \forall j$, giving us $\hat{p}$ for each product, and using $\hat{p}$ and the assumption that $\xi_j = \omega_j = 0, \forall j$ to calculate the derivatives needed for the last set of instruments. Reynaert and Verboven (2012, *working paper* "Improving the Performance of Random Coefficient Demand Models the Role of Optimal Instruments Preliminary Version") do this and a more detailed approximation. They assume that the (ξ, ω) have a distribution which is indpendent of (x, w) and normal, and compute a simulation estimator of the optimal instruments. That is they use the first stage estimates to parameterize the normal distribution, take draws from it, assume it is a unique price equilibrium and calculate it for each price, and then take averages over prices and the appropriate derivatives.[1] They report rather striking improvements from using their optimal instruments in a Monte Carlo study.

[1]They actually assume that the prices are perfectly competitive prices, and this simplifies the calculation considerably and insures uniqueness as prices simply equal marginal cost.

The Price Index Problem and the Interpretation of the Coefficients from Hedonic Regressions.

The problem of constructing a price index is a problem in welfare measurement. Here we show how looking at demand in characteristic space, can help with the issues that arise in constructing a price index. To do so, we will construct the hedonic pricing equation, and consider its properties. The discussion will clarify just what interpretation can be given to the coefficients estimated in the hedonic function.

For specificity we will consider the Consumer Price Index (the CPI) which is a measure of the "cost of living". We begin by providing a definition of the cost of living, and then consider some of the problems involved in estimating it. Similar problems arise in all price indices. So this should be a help in analyzing the issues that arise in both correcting the problems, and in using the price data.

The cost of living is defined as the cost of obtaining a given (usually a base year) level of utility, say $\overline{U}$ and, if we are looking at a household with characteristics $z_{i,t}$, this cost is calculated as

$$e_{i,t} = min_{q_{i,t}} \sum p_{j,t} q_{i,j,t}$$

subject to

$$U(q_i, z_{i,t}) = \bar{U}.$$

This implies

$$e = e(\text{choice set}_t, z_{i,t}, \bar{U}).$$

Note that the phrase "choice set" refers to the set of products available *and* their prices.

Consequently the "Laspeyres" true (or ideal) cost of living change between two periods compares the cost of attaining base period utility

$$CPI^T = \frac{e(\text{choice set}_1, z_{i,1}, \bar{U}_{i,0})}{e_{i,0}}.$$

The cost of living index *actually calculated* by the BLS is essentially

$$CPI^A = \frac{\Sigma_j w_j p_{j,1}}{\Sigma_j w_j p_{j,0}}$$

where the weights are determined by a previous consumer expenditure survey[2], and the prices are obtained largely from enumerators who sample randomly selected items within given expenditure groups (e.g. computers, T.V.'s) at randomly selected outlets.[3] They then return to re-sample the same goods after two months (sometimes one month). The price ratios obtained from the two periods sampled are averaged for the various component indices and then weighted with the CEX weights. This procedure generates what is called a matched model index[4].

If there were a single individual, and if all goods available in the base period are available in the given period, this index would insure that the consumer can purchase the same bundle of goods in the given as in the base period, and hence would insure that consumer the base periods' level of utility. This is

[2]From the CEX; which is a quarterly expenditure survey on a rolling panel of households

[3]A point of purchase survey is used to determine where to sample from.

[4]If that is all they did it would be a pure matched model index, however they are in the process of modifying it to account for a number of problems, in particular to the problem we are about to consider.

the (Konus)"rationale" for constructing the index as a matched model index. If there are many consumers, and we weighted accordingly, we would obtain a weighted average of the change in income that would insure that each individual could obtain the same goods they obtained yesterday.

There was a commision set up by the senate to evaluate the biases in the CPI (the Boskin Commission). They estimated two biases

- Substitution bias. This is the bias that arises from holding the basket of purchases fixed when the price vector changes. When that vector changes the consumer can do better by switching the quantities purchased. If one had estimates of the distribution of utilities one could estimate this bias directly. The commission, with some hand waving, came up with a number of .4% a year.

- New Goods biases in matched model indices. New goods are never evaluated directly. They do come in with sample rotation and replacement procedures, but then we only evaluate their price changes after they came in. So we never compare the new goods to the old goods, which is where part of the gains from new goods come. With even more hand waving, the Boskin commission set this at .7% per year).

Remedies

- Faster sample rotation procedures. Impact depends on the I.O. of good introduction; not studied much in economics. If the good is introduced at a low introductory price to attract consumers, and then rises when consumers become aware of it. Faster rotation could bias the index further. *Note:* eventually you want the new goods included; else you will be computing an index of goods people do not purchase. The only question is whether to introduce them soon after introduction or after prices have settled down.

- The selection problem and hedonic indexes. The selection problem is that the enumerators often do not find the same good on the shelf in the comparison period. If those goods are simply dropped, we incur a selection problem. The goods which are not on the shelves are disproportionately goods whose prices were falling. So we are dropping the observations from the left tail of the distribution of price changes and getting an average which is biased upwards. We can correct for this if we are willing to move to characteristic space by using hedonic indices; hold characteristics constant and compare prices for a given set of characteristics. This dates back to Court(1939) and Griliches(1961). Though this does not get at the inframarginal returns to new goods, it does allow for a correction for goods that exit the market (see the Tables below).

Use of hedonics in the CPI.

- Do a hedonic regression in every period. Price the goods that were available in $t-1$ but not in t at the hedonic prediction for their price in year t.

- Provided that a good with the same characteristics is available in both periods, the rationale for using this predicted price ratio is just the lower bound argument from Konus in characteristics space.

- The big benefit from using hedonics is that it partially controls for sample selection. Matched model indices throw out price comparisons for "characteristic bundles" that are withdrawn from the market and those are likely to be "characteristic bundles" whose market value has fallen (that is why they are withdrawn).

- Of course it only corrects based on observed characteristics. If the product has important characteristics that are not included in the data set being used, and it is the obsolescence of those characteristics that is determining the selection bias, we need a more complex correction procedure; one that, to the extent possible, corrects also for unobserved characteristics (see Erickson and Pakes, AER, 2011) .

A final note on this. There is no direct attempt to control for new goods biases here. It is easiest to see what those biases are if we assume that the (real) price of a new good is highest at the time the good is introduced, and falls thereafter (as it is obsoleted by new competing goods), which is often true for technology goods. Then I can cleanly partition the consumer surplus generated from the new good into two sources

- the inframarginal surplus obtained by the consumers who bought the new good in the "initial period", i.e. the period before the new good enters the index,

- and the gains that arise from the price falls after the new good enters the index.

The second source of gains are not limited to the fall in prices of the good in question, as tyically its entrance will cause price falls of competing goods. However all of these price falls will be picked up in the matched model index. Heurisitically we have a price at which the consumer did not buy the new good, and a price at which she did buy, and that allows us to get a fairly sharp measure of the consumer surplus gains.

The first source of gains is much harder to measure. To measure it we would need to estimate the distribution of utility functions. Moreover were we to attempt to do that non-parametrically in product space the data would tell us that it was not identified. This is because there are no prices to trace out the gain in utility from individuals who bought the good before the new good entered the index. Characteristic based demand models do a little better on this. I.e. provided the characteristics of the new good lie in the span of the characteristics of products bought at some other time (and this could be either before or after the new good entered), you might actually be able to value the new good in characteristic space. But for the BLS to do this would require them to estimate utility functions, and that probably will not happen for a while (if ever), as it will open them up to complaints about the assumptions that must be made in order to obtain such estimates.

Uses and Abuses of Hedonic Regressions.

There are important details which I skipped over in describing the hedonic selection correction procedure above, and it is worth spending some time on them. Hedonics are used in all sorts of settings (evaluating amenities, homes, clean air, lives...), as well as in constructing the CPI, and similar issues will arise in other contexts. Fundamentally, to understand what a hedonic regression represents, and how we can use it, we have to understand how prices are formed.

If we begin with a market in which all firms are single product firms playing Nash in prices and there are constant marginal costs given by $mc(\cdot)$, then as derived above prices are

$$p_i = mc(\cdot) + \frac{D_i(\cdot)}{|\partial D_i(\cdot)/\partial p|},$$

where the second term, $\frac{D_i(\cdot)}{|\partial D_i(\cdot)/\partial p|}$, is the mark-up which varies inversely with the elasticity of demand at the point. The hedonic function, say $h(x)$, is the expectation of price conditional on x.

$$h(x_i) \equiv E[p_i|\, x_i] = E\left(mc(\cdot)|\, x_i\right) + E\left(\frac{D_i(\cdot)}{|\partial D_i(\cdot)/\partial p|}|\, x_i\right),$$

where the expectation integrates over randomness in the processes generating the characteristics of competing products, input prices (which may also contain a markup), and productivity.

There are two important points to take from this equation. First the markups are determined by the goods competing in the market and the distribution of consumer tastes. As a result, they *should not generate stable functions of characteristics* either:

- across time in a given market, or
- across markets.

The implication is that if one is going to use the hedonic regression to predict prices, one has to recompute that regression equation in every period in every market. To see how important this is, consider the computer chip market. The high end computer chip today will earn a large markup over marginal cost (we will see this formally in our discussion of the vertical model below). However, as new faster chips come in tomorrow, the markup on today's high end chip will fall to zero, and eventually the chip will be retired (its variable profits will be less than its fixed costs).

The second lesson to learn is that the regression function coefficients have no interpretation and are of no use in and of themselves. These coefficients are a mix of the coefficients from the cost equation and the markup equation and since markups are a result of a complicated process, can have signs that are unintuitive.... All that matters for what we will do is the price prediction as a whole not individual coefficients.

Iain Cockburn (NBER, 1997) illustrates this point rather dramatically. The NIH did tests on two types of drugs to alleviate rheumatoid arthritis. Both drugs helped alleviate the symptons of arthritis by about the same extent but one had fairly serious side effects on most, though not all, patients. The one without the side effects worked on most patients but there was a subgroup on which it did not work. The probability of the second class of drugs, the drugs with the side effects, working was statistically independent of whether the first set of drugs worked on the patient. As a result, several drug companies rushed into developing drugs of the first type, and the markup on those drugs fell. In contrast the market for the second type of drug could only sustain one firm, and that firm charged a monopoly price. As a result, the hedonic regression, which included serious side effects as a right hand side variable, got a highly significant positive coefficient on that variable; the second group of drugs, the drugs that produced side effects sold at a higher price. This obviously should not be interpreted as patients preferring serious side effects.

The two points that this discussion is meant to stress are that

- Any procedure which requires the assumption that the hedonic function is stable across times or markets is likely to be problematic, and
- Any procedure which requires an intepretation of individual hedonic coefficients is likely to be problematic.

It is worth elaborating on the latter point. If one looks at the pricing equation it has two components; a cost function, and a markup term. The markup term will be relevant provided we do not abide by the fiction of price equals marginal cost (which would mean that any firm which has either fixed or sunk costs would go broke). So even if one were trying to estimate the cost function, we would have to account for the unobservable, and typical instrumental variables would not be valid. This because any variable which is related to the product's characteristics is likely to be related to the characteristics of other products in the market, and they determine the value of the markup. Moreover as illustrated above one can not even "sign" the relationship between the hedonic coefficients and "vertical" characteristics in utility models, never mind intepret their magnitudes.

Computer Example. Look to table 1. This shows the seriousness of the selection process, the fraction of computers sampled in the base period that were available for price quotes in the comparison period varied between 8 and 23%[5].

Look to table 2. Not only does the matched model index get the sign wrong, it also is negatively correlated with the hedonic. The reason is clear, years when a new generation of chips came in were years when a lot of old chips dropped out. The chips that remained were largely the earliest of the newest generation of chips, and their prices did not fall.

The BLS is now in the process of updating their data gathering procedure in a way that allows them to use hedonic indices based on regressions that are estimated separately in every period in "production mode". To go further than this we would need a demand system, and the BLS and other government agencies are hesitant to do this except in an "experimental" way for fear of engaging in judgement calls. Even if they were to do this, the current methodology is essentially assuming a separable utility function, which could be problematic.

[5] These are not the data the BLS uses to compute the computer component index, but the average drop-out rate in this data is virtually identical to the drop-out rate in

Table 1: Characteristics of Data*.

year	95	96	97	98	99
# observations	264	237	199	252	154
matched to $t+1$	44	54	16	29	n.r.
characteristics					
speed (MHz)					
min	25	25	33	140	180
mean	65	102	153	245	370
max	133	200	240	450	550
ram (MB)					
min	2	4	4	8	16
mean	7	12	18	42	73
max	32	64	64	128	128
hard disk (GB)					
min	.1	.1	.2	.9	2
mean	.5	1	1.8	4.5	8.5
max	1.6	4.3	4.3	16.8	25.5

Taken from Pakes(2003), *American Economic Review.*.

Note on Unobserved Characteristics and Selection. This is taken from Erickson and Pakes (*AER*,2011). The hedonic correction assumes that the selection is based on observable characteristics.

TV Example.

- there is 20% turnover over the sampling interval (almost identical rate to that in computers),
- there is ample evidence indicating that the goods that exit have prices that are falling disporportionately.

Yet when we compute a hedonic index based on a set of characteristics comparable to what the BLS analyst uses we get an index which is

- about the same value as the mm index (as noted in an NAS

Table 2: Alternative PC Price Indexes*

	Year	95/96	96/97	97/98	98/99	av.
1. Hedonics	base	-.097	-.108	-.295	-.155	-.164
		(.040)	(.063)	(.045)	(.099)	(.091)
(or, I)	f.a.	-.094	-.111	-.270	-.150	-.156
		(.039)	(.052)	(.044)	(.054)	(.079)
2. Matched	Tornquist	.012	.002	.09	.011	.028
model (M)	Laspeyres	-.013	-.002	-.08	-.011	-.027
	percent matched	16.6	22.8	8.0	11.5	14.7
6. Dummy	base	-.135	-.098	-.160	-.170	-.141
		(.038)	(.035)	(.027)	(.040)	(.032)
Variables	f.a.	-.152	-.122	-.213	-.143	-.158
		(.040)	(.032)	(.041)	(.028)	(.039)

*Taken from Pakes (2003), *American Economic Review*. Standard errors appear in brackets below estimate. They are estimated by a bootstrap based on 100 repetitions. "base" refers to base specification and "f.a." refers to fully loaded specification in table 2. "n.r." = not relevant.

volume on reforming the price indices),

Annual Computer vs Bimonthly TV Data: Unobserved Characteristics and Selection Corrections.

- Most of our TV characteristics are dummy variables indicating the presence or absence of advanced features.
- Exit is disproportionately of high priced goods that have most of these features. They exit because they are obsoleted by newer high priced goods with higher quality versions of the same features.
- There are no cardinal measures for the quality of these features.
- As a result, in the TV market selection is partly based on characteristics the analysts can not condition on, i.e. on what an econometrician would call "unobservables".
- The hedonic controls for changes in values of observed characteristics of all goods, but misses changes in value of unobserved characteristics. The matched model controls for changes in value of unobserved and observed characteristics of continuing goods, but does not control at all for exiting goods.
- Look to regressions for observable characteristics.
- Look at residuals for about to exit versus continuing goods.
- Develop procedure which accounts for unobserved characteristics (and maintains the bound, and is robust to assumptions and data sets). The procedure requires the fea-

tures of the data and the selection process. In particular, when doing the hedonic we condition on unobservable characteristic in previous period (and where available price changes in the interim). It uses only the original small number of characteristics (which got an 8.8% fall initially); we get the next table.

Table 1: **TV Indices.**

Index Calculated	MM	hedonic
Using Log-Price Regression For Hedonic		
hedonic uses S24	-10.11	-10.21
s.d. (across months)	5.35	7.53
S24 % l.t. mm^2		.50
hedonic uses S9	-10.11	-8.82
s.d. (across months)	5.35	7.05
S9^4 % l.t. mm		.40
Taken from Ericson and Pakes (forthcoming) *American Economic Review.*		

Table 2: **Hedonic Regressions: Dependent Variable is Log-Price**

Regressors	mean R^2	mean adj R^2	min R^2	min adj R^2	max R^2	max adj R^2
S5	.896	.894	.873	.870	.913	.911
S10	.956	.954	.942	.937	.967	.965
S24	.971	.967	.959	.953	.978	.975

R^2 statistics from log-price regressions run on each month from March 2000 to January 2003.

Market Level Paper.

- The "BLP" instruments (which is unfair to us) trace out substitution patterns (in the BLP model context those correspond to getting the θ on the random coefficients).

Table 3: **Hedonic Disturbances for About to Exit, Recently Entered, Goods.**

Variable	All Continuing	a-Exit	r-New	Remaining Goods.
Using the S10 Specification for the Hedonic Regression[1].				
mean	-.0027	-.0157	-.0049	-.0022
s.d. of mean	.0010	.0026	.0021	.0014
s.d.(across months)	.0090	.0150	.0130	.0130
percent < 0	.6207	.8966	.5517	.5172

[1] See the description of the S10 specification.

Table 4: **Alternative Monthly Indexes for TV[1].**

Index Calculated	matched model	hedonic	hedNP
Bimonthly Data Only. Panel B: Non-Parametric Selection Model.			
mean	-10.11	-11.17	n.c.
standard deviation	5.35	5.01	n.c.
%l.t.mm		.80	n.c.
Using Monthly Data.			
Panel C: Probabilities and Price Changes.			
mean	-10.11	-12.51	n.c.
standard deviation	5.35	7.94	n.c.
%l.t.mm	n.c.	.83	n.c.

- The cost side instruments are particularly valuable for getting a separate effect of price.
- Non-parametrically you need $2J$ instruments, for the J shares and J prices.
- If you add a supply side model you find particular functions of the (x, p) which if we form a $\hat{p}$ which is a function of exogenous variables, should help on p in a way that uses the other $x's$ (BLP,99). One thing thats of interest is that the model doesn't need to be exactly correct, and the intuition goes through. See Verboten($\cdot$).

Micro Level Data.

- For a given level of demand you can get all effects of observed demographics from the micro data. You are left with the unobserved demographics, and the effect on the market level constant.

- To get the unobserved demographics and know pricing equation, you are relying on linearity (or maybe monotonicity)of the utility in the observed demographic and independence of the distribution of the random coefficient from the observable z.

- IF we were to add a pricing equation we could go further on this.

- The constant terms. You have cost shifters, you have the multi-product firm aspect, and then any functional form restriction on the utility function will generate for you functions of otehr $x's$.

- If the data contains different markets we go back to the market level data issues. The difference here is we can use differences in distributional aspects of the demographics as instruments.