Appendices to
Slutsky meets Marschak
The First-Order Identification of Multi-product Production

E. Glen Weyl

December 2009

THE SECOND OF THESE APPENDICES IS PRELIMINARY AND INCOMPLETE.

I. Marschak

A. Stochastic interpretations of the non-stochastic model

In first-order identification, I suppress all latent variables. This is problematic because the levels and derivatives which the model focuses on identifying may depend on the values of these latent variables. A simple example would be if demand were $Q(P, \epsilon) = a - b\epsilon P$ in which case the elasticity of demand measured would depend critically on the value of the latent ϵ. Nonetheless results derived from a model neglecting such latent variables can be informative even when they exist so long as it is interpreted properly and appropriate assumptions are made. Here will discuss three consistent interpretations of the non-stochastic model. All of these make strong assumptions which may be implausible in many applications. However I believe both that these assumptions will likely be relaxed by future research and that corrections for failure of these assumptions can likely be corrected in construction of concrete estimators, if first-order identification succeeds, in a way that is unlikely to be possible if first-order identification indicates a failure of identification.

First and most simply, suppose that the latent variables have a sufficiently small range over which the derivatives and levels of the structural function do not change substantially. Such latent variables are still a source of statistical noise, as $\mathbf{X}$ itself only varies over a limited range, but do not create large enough effects to change the relevant derivatives or function of these derivatives. For example in $Q(P, \epsilon) = a - b\epsilon P$ if ϵ is always close to 1 we know that the effect of an increase in P on Q will always be close to b; ϵ makes it difficult to identify b if P only varies a limited amount as well, but it does not interfere with the interpretation of b or bias its estimation. Given the implausibility (see below) of the linear separability assumptions invoked in many models, I suspect that the intuitive basis of such assumptions is actually beliefs that latent variables have a limited range. This is much like the reason why, along with second-order terms, interaction terms are often left out of linear regressions. Thus, despite its seeming specialness, I think in many ways this is the most robust justification of the approach I take.

In some cases, latent variables may be separable in the appropriate way from the model. While such assumptions are common, the appropriate sense of separability is quite sensitive to the derivative for function of derivatives that need to be estimated. For examples if the derivative of demand is the estimand then under linear separability $Q(P, \epsilon) = Q(P) + \epsilon$ variability in ϵ will not bias or muddy the interpretation of our estimate as long as P is independent of ϵ. However if the estimand is the pass-through rate $\frac{1}{2-\frac{QQ''}{(Q')^2}}$ then the necessary separability is that of the log of Q. That is the desired form is $Q(P, \epsilon) = \epsilon Q(P)$. If, on the other hand, we have linear separability, the pass-through rate ϵ and measuring, say, the mean value of Q and how this varies (to the first and second-order) with P will be insufficient to estimate a mean pass-through rate. In general, difficult-to-estimate properties of the joint distribution of Q' and Q would need to be observed. There are certainly cases when the appropriate form of separability is a reasonable assumption, but they seem uncommon.

Third, if the quantity of interest is linear in the observable derivatives, then the model may be interpreted as an average effect which may be what is of interest. For example, Harberger's formula states that welfare is $t\frac{dA}{dt}$ so if one's interest is in expected welfare only the average value of $\frac{dA}{dt}$ is of interest. Because $E\left[\frac{\partial \mathbf{f}}{\partial \mathbf{X}}\right] = \frac{\partial \mathbf{F}}{\partial X}$ the non-stochastic model can just be viewed as the average of the stochastic model. However if the estimand is non-linear in the observable derivatives, this is clearly invalid: $E\left[\frac{1}{2-\frac{QQ''}{(Q')^2}}\right] \neq \frac{1}{2-\frac{E[Q]E[Q'']}{(E[Q'])^2}}$.

Thus the set of cases where first-order identification is *directly* empirically relevant is limited, probably to a minority of cases of interest. However I still think it is *relevant*, if a bit more indirectly. It seems clear that measuring pass-through rates based on variation in prices requires some sort of measurement of the second-order effect of price variation on demand, even if this must be done in a more sophisticated way for general latent variable structures. Thus the non-stochastic first-order identification result tells a smaller part of the story in these cases, but it still seems a good, simple transparent starting point for the richer econometric analysis needed to complete local identification and, eventually, estimation. Clearly econometric work clarifying the amount of data needed to estimate non-linear functions of local derivatives under general structures for latent variables is important for making first-order identification practical for empirical work.

B. Extensions of the definition of first-order identification

In Subsection III.C I defined first-order identification for the case when all variables are finite dimensional and variation is around a single equilibrium point. Here I extend this definition in two directions.

First, consider the case in which the variables are again finite dimensional, but there are $M > 1$ equilibrium points about which variation is observed.

Definition 1′. *The* equilibrium levels *of the structure* $\mathbf{L}$ *is a* $M\left(|\mathbf{X}|+|\mathbf{Y}|\right)+|\mathbf{U}|$*-dimensional real vector* $\left(\mathbf{X}_1^\star,\dots,\mathbf{X}_M^\star,\mathbf{U}^\star,\mathbf{F}\left[\mathbf{X}_1^\star,\mathbf{U}^\star\right],\dots,\mathbf{F}\left[\mathbf{X}_M^\star,\mathbf{U}^\star\right]\right)^\top$.

Definition 2′. *Let* i *be a natural number. The* i-th order varying derivatives*, denoted by* $\mathbf{D}^i$*, of the structure is a* $M|\mathbf{Y}|\begin{pmatrix}|\mathbf{X}|+i-1\\ i\end{pmatrix}$*-dimensional real column vector.* $\mathbf{D}^i$ *represents the* i*th order partial derivatives*[1] *of entries of* $\mathbf{Y}$ *with respect to entries of* $\mathbf{X}$ *at each point of the equilibrium points.*

Definition 3′-7′ are identical to Definition 3-7 respectively except that the single point $\mathbf{X}^\star$ is replaced by the vector $\left(\mathbf{X}_1^\star,\dots,\mathbf{X}_M^\star\right)^\top$.

Second, I consider the case when $\mathbf{X}$ is function on $\mathbb{R}$ and $\mathbf{f}$ and $\mathbf{g}$ are functionals. For simplicity I allow only a single equilibrium function $X^\star(\cdot)$. A classic example is Saez (2001)'s analysis of non-linear taxes, which are a schedule as a function of income. Of course this could be extended to the case of schedules on more general spaces, but this example should suffice to show how the basic approach can be extended. In fact I will focus specifically on the tax application, assuming $\mathbf{X}$ is a piece-wise linear function, possibly with an infinite but countable number of kinks.

Definition 1″. *The* equilibrium levels *of the structure* $\mathbf{L}=\left(\tilde{\mathbf{L}},X^\star[\cdot]\right)$ *where* $\tilde{\mathbf{L}}$ *is a* $|\mathbf{Y}|+|\mathbf{U}|$*-dimensional real vector* $\left(\mathbf{U}^\star,\mathbf{F}[\mathbf{X}^\star,\mathbf{U}^\star]\right)^\top$ *and* $X^\star$ *is a function mapping* $\mathbb{R}$ *to* $\mathbb{R}$.

Definition 2″. *Let* i *be a natural number. The* i-th order variations *of the structure* $\mathbf{D}^i$ *is a function from* $\mathbb{R}^i$ *to* $\mathbb{R}^{|\mathbf{Y}|}$ *with the property that* $\mathbf{D}^i$ *is permutation invariant in the sense that* $\mathbf{D}^i(z_1,\dots,z_i)=\mathbf{D}^i(z_{j_1},\dots,z_{j_i})$ *where* $\{j_k\}_{k=1}^i$ *is any permutation of the naturals from* 1 *to* i.

[1] For concreteness and without loss of generality, assume these are listed in lexicographic order first in the order of the equilibrium points, second in the ordering of $\mathbf{Y}$ and then in the monomial lexicographic ordering with respect to the entries of $\mathbf{X}$.

The jth entry of $\mathbf{D}^i$, $D^i_j(z_1, \dots z_i)$, represents the ith-order variation of Y_j with in the directions of the piecewise functions

$$\tilde{X}_{z_j}(z) = \begin{cases} X^\star(z) & z \leq z_j \\ X^\star(z) + z - z_j & z > z_j \end{cases}$$

which form a complete basis for space of piecewise linear functions from the reals to the reals.

Note that the basis here corresponds to Saez (2001)'s perturbation in the direction of uniformly increased marginal tax rates above some threshold income. Definitions 3″-5″ are identical to definitions 3-5 except that the relevant spaces are changes and "function" is replaced by "functional".

Definition 6″. *A property of $(\mathbf{F}, \mathbf{G})$ at some point $\left(\tilde{X}[\cdot], \tilde{\mathbf{U}}\right)$ is said to be* point identified *given an N-th order observability condition Θ^N at point $(X^\star[\cdot], \mathbf{U}^\star)$ and a structural restriction S if $\forall \mathbf{x} \in \mathbb{R}^{\left|\Theta^N_\cdot\right|}$ (or functions if Θ^N is infinite-dimensional) every functional $(\mathbf{F}, \mathbf{G}) \in S$ satisfying $\Theta^N_\cdot = \mathbf{x}$ and $\Theta^N_> > 0$ at that point assigns the same value to the property of interest (as a function of $\mathbf{x}$). A property is said to be* sign identified *if, over the same range of settings, the property is always of the same sign (positive, negative or zero). The property will be said to be* point unidentified *or* sign unidentified *whenever they are not point identified or sign identified respectively.*

Definition 7″. *An N-th order observability condition Θ^N at point $(X^\star, \mathbf{U}^\star)$* over-identifies *a structural restriction S if $\exists \mathbf{x} \in \mathbb{R}^{\left|\Theta^N_\cdot\right|} : \exists (\mathbf{F}, \mathbf{G}) \in S$ satisfying $\Theta^N_\cdot = \mathbf{x}$ and $\Theta^N_> > 0$ at that point* and *$\exists \mathbf{x}' \neq \mathbf{x} : \nexists (\mathbf{F}, \mathbf{G}) \in S$ satisfying $\Theta^N_\cdot = \mathbf{x}$ and $\Theta^N_> > 0$ at that point. In the case when Θ^N is infinite dimensional these are replaced with the corresponding functional equivalents.*

C. Proof

Proof of Proposition 1. First I establish the positive result. By the chain and multiplication rules

$$\frac{do}{dt} = W'P' + Q + tQ'P'$$

By Cournot's formula $P' = \frac{1}{2 - \frac{Q''Q}{(Q')^2}}$ and by Hotelling's lemma $W' = -Q$. We are evaluating at $t = 0$, so

$$\frac{do}{dt} = Q\left(1 - \frac{1}{2 - \frac{Q''Q}{[Q']^2}}\right) = \frac{Q\left([Q']^2 - Q''Q\right)}{2\,(Q')^2 - Q''Q}$$

By the chain rule $Q_x = Q' P_x$ and by the multiplication and chain rules $Q_{xx} = Q'' (P_x)^2 + Q' P_{xx}$. Weyl and Fabinger show that declining marginal revenue is equivalent to $2 (Q')^2 > Q'' Q$ and $Q > 0$ so the sign of $\frac{do}{dt}$ is the same as that of

$$(Q')^2 - Q''Q = \left(\frac{Q_x}{P_x}\right)^2 - \frac{Q\left(Q_{xx} - \frac{Q_x}{P_x} P_{xx}\right)}{(P_x)^2} = \frac{(Q_x)^2 P_x - Q (Q_{xx} P_x - Q_x P_{xx})}{(P_x)^3}$$

Recall that $P_x > 0$ if any function obeys the observability conditions in S. Thus the sign of $\frac{do}{dt}$ must be negative; a subsidy is desirable.

To establish point unidentification, note that the observables imply a local level and elasticity for demand, but only restrict the pass-through to be greater than 1. It is well-known that for any local elasticity and level of demand, pass-through for a decreasing, declining marginal revenue demand maybe anything from 0 to ∞; Bulow and Pfleiderer (1983) show how to make the explicit correspondence. Letting the pass-through be ρ we know from above that $\frac{do}{dt} = Q(1 - \rho)$. If $\rho = 1.5$ this clearly does not yield the same value for $\frac{do}{dt}$ as if $\rho = 2$, but both obey the observability conditions and have declining marginal revenue if we choose them from the Bulow-Pfleiderer class. Thus $\frac{do}{dt}$ is point unidentified.

D. An example of the flexibility of first-order identification

In merger analysis two leading approaches correspond to first-order and parametric identification: the Werden-Froeb-Tschantz-Farrell-Shapiro local test discussed above and the Werden and Foreb (1994)-Nevo (2000) merger simulation approach respectively. Baker and Reitman (Forthcoming) argue that while the former method is simpler, more transparent and more robust, the second method may be allows higher order effects (through own or other firm prices) to be explicitly taken into account as a full, post-merger equilibrium is solved for, an argument also emphasized by Pakes (1997). This argument is somewhat misleading, however. The Werden-Farrell-Shapiro framework could easily be extended to include second-order effects as suggested by Froeb et al. (2005). Identifying these would require either additional assumptions or extra data. Merger simulation ties these down through functional form assumptions which either directly impose their magnitudes or tie these to some first-order-relevant magnitude, again by assumption. If those assumptions are plausible, they can be made explicitly as part of the second-order identification analysis, reducing the data burden there; if they are implausible then any extra identification provided by the parametric approach is unpersuasive. Thus despite the apparent incompleteness of first-order identification analysis, it can easily be extended to provide higher order approximations, making use of as many assumptions as are plausible to impose and as much data as is available.

II. Slutsky

A. Second-order conditions for oligopoly

In the case of monopoly and perfect competition, simple conditions for a unique solution to the equations involve forms of concavity. Under oligopoly the firm's choices do not maximize a coherent objective function and thus the object on which to impose concavity is unclear. However an old literature from general equilibrium theory (Gale and Nikaido, 1965) gives conditions for unique solutions to systems of equations. The natural extension to Cournot oligopoly is that the gradient of marginal revenue be everywhere a matrix, the negative which has all positive principal minors. Weaker conditions are also possible (Kolstad and Mathiessen, 1987), but not if the equilibrium is to be unique for all cost vectors. In the paper I just assert a unique behavior of all other firms as a function of the firm of interest's behavior. To justify this would require applying these conditions to the joint Jacobian of all firms' profit functions.

B. Parametric and non-parametric identification

Here I consider the parametric and non-parametric identification of the multi-product producer models. I only consider the nonparametric argument, as the parametric one is a special case. First note that, so long as the cost shifts induce arbitrary range shifts in quantities, which they must if they can become independently arbitrarily large or small (given differentiability and therefore finiteness of derivatives of cost), they trace out the full demand surface. Second note that, given this, the vector of marginal costs can be recovered at each quantity by Rosse (1970)'s argument applied to the multi-product pricing rule of Ramsey (1927), as observing demand makes both marginal revenues observable. This supplies a differential equation that defines cost, up to an irrelevant additive constant, at every quantity pair. This argument implies full identification of the parametric model which has no such additive constant and identification of the nonparametric model up to a fixed cost.

C. Proofs

Proof of Theorem 1. I establish each piece of the theorem in turn:

1. If the structural functions are twice differentiable with respect to $\mathbf{X}$ then so must profits be, as $\mathbf{X}$ enters profits only through the observable $\mathbf{c}$. But by Milgrom and Segal (2002)'s general envelope theorem (Theorem 1), given that under any of perfect competition, monopoly or consistent conjectures oligopoly $\mathbf{Q}$ maximizes a residual profit function $\pi(\mathbf{Q};\mathbf{x})$, $\frac{d\pi}{dx_i} = \frac{\partial \pi}{\partial x_i} = -\frac{\partial c_i}{\partial x_i}$. By the hypothesis of twice differentiability and the chain rule
$$\frac{d^2 c_i}{dx_i dx_j} = \frac{d\frac{\partial c_i}{\partial x_i}}{dx_j} = \frac{\partial^2 c_i}{\partial x_i \partial Q_i}\frac{\partial Q_i}{\partial x_j}$$
Identical reasoning for commodity j yields
$$\frac{d^2 c_i}{dx_i dx_j} = \frac{\partial^2 c_j}{\partial x_j \partial Q_j}\frac{\partial Q_j}{\partial x_i}$$
and thus
$$\frac{\partial Q_j}{\partial x_i} = \frac{\frac{\partial^2 c_i}{\partial x_i \partial Q_i}}{\frac{\partial^2 c_j}{\partial x_j \partial Q_j}}\frac{\partial Q_i}{\partial x_j}$$
This establishes the point identification result and the assumption that both terms in the quotient on the right-hand side are positive establishes sign-identification.

2. By identical logic to above, under perfect competition, $\frac{d\pi}{dP_i} = Q_i$ and thus $\frac{d^2\pi}{dP_i dx_j} = \frac{\partial Q_i}{\partial x_j}$. But $\frac{d\pi}{dx_j} = -\frac{\partial c_j}{\partial x_j}$ so $\frac{d^2\pi}{dP_i dx_j} = -\frac{\partial^2 c_j}{\partial x_j \partial Q_j}\frac{\partial Q_j}{\partial P_i}$ Thus
$$\frac{\partial Q_j}{\partial P_i} = -\frac{\frac{\partial Q_i}{\partial x_j}}{\frac{\partial^2 c_j}{\partial x_j \partial Q_j}} = -\frac{\frac{\partial Q_j}{\partial x_i}}{\frac{\partial^2 c_i}{\partial x_i \partial Q_i}} = \frac{\partial Q_i}{\partial P_j}$$
where the last two equalities follow from the results above and establish both the sign and point identification.

3. Coming soon.

4. Coming soon.

Proof of Theorem 2. I have assume that profits π is smooth as I have assumed all of its components are smooth. Thus if we let $\tilde{\mathbf{H}}$ be the matrix $\left[\frac{\partial^2 \pi}{\partial Q_i \partial Q_j}\right]_{i,j=1}^{N}$ this must be negative semidefinite at the equilibrium point under any of the (consistently) optimizing models as this is a necessary condition for optimization. Furthermore, by the envelope theorem and/or the necessary conditions for optimization letting $\mathbf{J} \equiv \left(\frac{\partial \pi}{\partial Q_1}, \ldots, \frac{\partial \pi}{\partial Q_N}\right)^\top$, at the equilibrium point

$\mathbf{J} = 0$. So letting $\mathbf{c}_{12}$ be the NxN diagonal matrix with entries $\frac{\partial^2 \pi}{\partial c_i \partial Q_i}$ by twice continuous differentiability

$$\nabla \mathbf{J}\mathbf{H} = \tilde{\mathbf{H}}\mathbf{H} = \mathbf{c}_{12}$$

thus $\tilde{\mathbf{H}}$ must be non-singular as $\mathbf{c}_{12}$ is a strictly positive diagonal matrix and so

$$\mathbf{H} = \tilde{\mathbf{H}}^{-1}\mathbf{c}_{12}$$

$\tilde{\mathbf{H}}^{-1}$ is negative definite as it is the inverse of a negative definite matrix and thus so is $\tilde{\mathbf{H}}^{-1}\mathbf{c}_{12}$ as negative definiteness is preserved under multiplication by a strictly positive diagonal matrix. Thus $\mathbf{H}$ is negative definite establishing the desired result.

D. Other general results

Over-identification results for symmetry are coming soon.

Perhaps the most famous identification strategy in empirical industrial organization and auction theory is the use of demand conditions to tie down a firm's marginal cost of production through her first-order condition. This approach, due to Rosse (1970), requires exogenous price variation (or direct observation of demand patterns in auctions). While not formalized as such, this is essentially a first-order identification result linking the elasticity of demand to firm marginal costs. One natural way to generate the necessary price variation is through cost shocks, as discussed throughout this section. To emphasize a basic principle above, that consistency of conjectures greatly reduces the dimensionality of the necessary observations, I now consider first-order identification of marginal cost levels through cost shocks (exogenous, product-specific price variation). The main message is that under perfect competition, monopoly and *only* consistent conjectures oligopoly marginal cost levels are revealed by marginal cost shocks *only to the relevant firm.* With inconsistent conjectures, marginal cost identification typically requires variation in all competitor's costs as well. For simplicity, I discuss the results for the case of a single product. Extending the results to the case of multi-product firms above would require cost shocks on all the products of the focal firm.

For this purpose, I add an unobservable endogenous cost variable $U = C(Q) + c(Q, X)$.

Theorem 1. *Consider the further restriction that, at $(x^\star, \mathbf{P}^\star)$ or $(x^\star, \overline{\mathbf{P}}^\star)$ as appropriate, the structural function is differentiable and suppose that $N = 1$.*

1. *Under perfect competition if the equilibrium price is observable, $\frac{\partial U}{\partial Q}$ is point identified*

2. *Under monopoly or consistent conjectures oligopoly if* $\frac{\partial P}{\partial x}$ *and* $\frac{\partial Q}{\partial x}$ *are point observable,* $\frac{\partial U}{\partial Q}$ *is point identified.*

3. *Under inconsistent conjectures oligopoly, even if* $\frac{\partial P}{\partial x}$ *and* $\frac{\partial Q}{\partial x}$ *are point observable,* $\frac{\partial U}{\partial Q}$ *is point unidentified.*

This formalizes the notion, hinted at by Baker and Bresnahan (1988), that their identification strategy of looking at single-firm cost shocks to recover marginal costs in an oligopoly is valid if and only if conjectures are assumed consistent. Thus consistency of conjectures has a broad identifying power in oligopoly models, distinct to other solution concepts.

Proof. Coming soon.

If conjectures are inconsistent then the "real world" elasticities observed when costs and prices respond to a cost shock are not those relevant to the firm's optimization. If instead of having consistent conjectures, the firms act as if the prices or quantities of their rivals were fixed, then variation in rivals prices or quantities will be needed to "cancel out" the strategic price and quantity adjustment induced by the cost change and thus cleanly recover the relevant demand elasticity. However under consistent conjectures the cost-induced effects are identical to those the firm takes into account when changing her choice variables and thus it is the "real world" elasticities that are relevant. This conundrum also arises in merger analysis (Farrell and Shapiro, 2010) and is one reason why consistent conjectures may be a useful, if under utilized, solution concept there.

E. Applications

Coming soon.

References

Baker, Jonathan B. and David Reitman, "Research Topics in Unilateral Effects Analysis," in Einer Richard Elhauge, ed., *Research Handbook on the Economics of Antitrust Law*, Northampton, MA: Edward Elgar, Forthcoming.

_ **and Timothy F. Bresnahan**, "Estimating the Residual Demand Curve Facing a Single Firm," *International Journal of Industrial Organization*, 1988, *6* (3), 283–300.

Bulow, Jeremy I. and Paul Pfleiderer, "A Note on the Effect of Cost Changes on Prices," *Journal of Political Economy*, 1983, *91* (1), 182–185.

Farrell, Joseph and Carl Shapiro, "Cannibalization, Pass-Through, and Market Definition," 2010. This paper is currently being prepared, but is cited and described by the authors in their 2008 paper cited here.

Froeb, Luke, Steven Tschantz, and Gregory J. Werden, "Pass-Through Rates and the Price Effects of Mergers," *International Journal of Industrial Organization*, 2005, *23* (9–10), 703–715.

Gale, David and Hukukane Nikaido, "The Jacobian Matrix and Global Univalence of Mappings," *Mathematische Annalen*, 1965, *159* (2), 81–93.

Kolstad, Charles D. and Lars Mathiessen, "Necessary and Sufficient Conditions for Uniqueness of Cournot Equilibrium," *Review of Economic Studies*, 1987, *54* (4), 681–690.

Milgrom, Paul and Ilya Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 2002, *70* (2), 583–601.

Nevo, Aviv, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 2000, *31* (3), 395–421.

Pakes, Ariel, "New Economic Tools for Merger Analysis," 1997. http://www.economics.harvard.edu/faculty/pakes/files/mslides.pdf.

Ramsey, Frank P., "A Contribution to the Theory of Taxation," *Economic Journal*, 1927, *37* (145), 47–61.

Rosse, James N., "Estimating Cost Function Parameters Without Using Cost Data: Illustrated Methodology," *Econometrica*, 1970, *38* (2), 256–275.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205–229.

Werden, Gregory J. and Luke M. Foreb, "The Effects of Mergers in Differnentiated Porducts Industries: Logit Demand and Merger Policy," *Journal of Law, Economics and Organizations*, 1994, *10* (2), 407–425.