A Comment on "Simple Estimators for Invertible Index Models"

Andres Aradillas-Lopez

Pennsylvania State University

August 5, 2016

1 Introduction

The approach described in Ahn, Ichimura, Powell, and Ruud (2015) (henceforth AIPR) constitutes an elegant contribution to the "control function" literature, which focuses on semiparametric models where endogeneity or nonlinearities of unknown form can be captured by a control function (or "control variable"). Like the rest of this literature, the applicability of the method in AIPR presupposes the ability to identify (perhaps nonparametrically) the control function in order to have the ability to "match" (asymptotically) pairs of observations in a way that identifies the parameters of interest (see Equations (2.15)-(2.19) in AIPR). While the control function approach has been shown to have aide applicability (Ahn and Powell (1993), Honoré and Powell (1994), Blundell and Powell (2004), Honoré and Powell (2005), Imbens and Newey (2009), Hong and Shum (2009), Aradillas-Lopez, Honoré, and Powell (2007)), in this note I argue that it is easy to construct examples of microeconometric models where matching is not possible, either because of interval data, missing data or incomplete models (e.g, structural models with "multiple equilibria"). In doing this, I also want to argue that, while matching is not possible and the methodology in AIPR cannot be applied, the examples I present produce (conditional) moment inequalities which can be used to do inference on the parameters of interest by using recently developed methods involving conditional moment inequalities. This shows that control function models are powerful vehicles for inference even in partially identified settings.

2 A generic setup

For brevity, consider a model described as follows,

$$Y_i = X_i'\beta_0 + \phi_i + \nu_i. \tag{1}$$

In what follows ν_i is an unobserved latent variable that satisfies some basic exogeneity restriction[1] (e.g, mean-independence with respect to X_i) and ϕ_i is the *control function* which captures all the endogeneity of the model. The setting of AIPR (and that of the existing literature on control functions) assumes the existence of a known (or estimable) function g, which depends on a vector of observable covariates W_i and (possibly) an unknown parameter

[1] The exact condition will be clarified below.

γ (whose true value is denoted by γ_0) such that, for any pair of observations i, j,

$$g(W_i;\gamma_0) = g(W_j;\gamma_0) \quad \Longrightarrow \quad \phi_i = \phi_j \tag{2}$$

Therefore, in a manner analogous to panel data fixed-effects, matching $g(W_i;\gamma_0)$ and $g(W_j;\gamma_0)$ for any pair of observations allows us to difference out[2] ϕ_i and ϕ_j,

$$g(W_i;\gamma_0) = g(W_j;\gamma_0) \quad \Longrightarrow \quad Y_i - Y_j = (X_i - X_j)'\,\beta_0 + \nu_i - \nu_j$$

Inference on the model's parameters can proceed from here using, e.g, mean-independence conditions of the form $E\left[\nu_i|X_i, W_i\right] = 0$. Note that (2) is crucial for this approach to work.

2.1 A case where the matching conditon (2) does not hold but there are bounds for the control function

In this note I focus on a case in which (2) does not hold and the type of matching used in AIPR (and all the related literature) is impossible. To be precise, I focus in a case where (2) is replaced with the following condition[3],

$$\underline{g}(W_i;\gamma_0) \geq \overline{g}(W_j;\gamma_0) \quad \Longrightarrow \quad \phi_i \geq \phi_j \tag{2'}$$

where, for any given γ, both $\underline{g}(W_i;\gamma)$ and $\overline{g}(W_i;\gamma)$ are wither known or estimable. We mentioned above that, in these types of models, the latent variable ν_i satisfies some exogeneity restriction; for the sake of exposition let us assume that

$$E\left[\nu_i|X_i, W_i\right] = 0 \tag{3}$$

Together, the conditions in (1), (2) and (3) fit the framework of AIPR perfectly. My goal here is to argue that their results and ideas, combined with recent advances in moment-inequality models, can also help guide inference if we replace (2) with (2').

2.2 Examples

Specific examples of models described by (1) and (2') are not hard to construct. They can arise, e.g, in the context of interval data or incomplete (i.e, partially identified) economic models.

A sample selection model with interval data

Estimation with interval data has been studied, for example, in Manski and Tamer (2002), but it has not received attention in the context of control functions. Consider a basic sample selection model of the type studied by

[2] This is the essence of Equation (2.17) in AIPR.

[3] For everything that follows, (2') can be replaced with the following weaker condition,

$$\underline{g}(W_i;\gamma_0) \geq \overline{g}(W_j;\gamma_0) \quad \Longrightarrow \quad E\left[\phi_i|X_i, W_i\right] \geq E\left[\phi_j|X_j, W_j\right].$$

Heckman (1979), described as

$$Y_i^* = X_i'\beta_0 + \eta_i, \quad d_i = \mathbb{1}\left\{Z_i'\gamma_0 + \varepsilon_i \geq 0\right\}, \quad Y_i = \begin{cases} Y_i^* & \text{if } d_i = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Therefore $E\left[Y_i|X_i, d_i = 1\right] = X_i'\beta_0 + E\left[\eta_i|\varepsilon_i \geq -Z_i'\gamma_0\right]$. Suppose $(\eta_i, \varepsilon_i) \perp (X_i, Z_i)$, then for uncensored observations Y_i we can express

$$Y_i = X_i'\beta_0 + \mu(Z_i'\gamma_0) + \nu_i$$

Suppose that the joint distribution of ε_i and η_i is assumed to be such that $\mu(\cdot)$ is *nonincreasing*. A special case is bivariate Normality with positive correlation. This model is compatible with our generic description with $\phi_i = \mu(Z_i'\gamma_0)$. If Z_i is observable along with d_i, then methods like the ones proposed in AIPR are readily applicable (see Section 5 in AIPR). Suppose however that some elements in Z_i are not observable, but we observe W_i^L and W_i^U such that

$$W_i^{L\prime}\gamma_0 \leq Z_i'\gamma_0 \leq W_i^{U\prime}\gamma_0 \quad \text{w.p.1.}$$

In general, this assumption presupposes knowledge about the signs of at least some of the coefficients γ_0. Given the assumption that $\mu(\cdot)$ is nonincreasing,

$$W_i^{U\prime}\gamma_0 \leq W_j^{L\prime}\gamma_0 \implies Z_i'\gamma_0 \leq Z_j'\gamma_0 \implies \phi_i \geq \phi_j.$$

Therefore this model fits the condition in (2') if we define $\underline{g}(W_i;\gamma_0) = -W_i^{U\prime}\gamma_0$ and $\overline{g}(W_i;\gamma_0) = -W_i^{L\prime}\gamma_0$.

A partially linear model with missing data

Microeconometric models involving agents' expectations arise in the context of many economic models. Consider one described as follows,

$$Y_i = X_i'\beta_0 + F(E_i[\xi_i|W_i]) + \nu_i, \tag{5}$$

where $F(\cdot)$ is an unknown but nondecreasing function. $E_i[\xi_i|W_i]$ denotes individual i's expectation of some outcome ξ_i conditional on W_i. If a random sample of (ξ_i, W_i) were observed and rational expectations were assumed, such that $E_i[\xi_i|W_i] = E[\xi_i|W_i] \ \forall\, i$, then expectations $E_i[\xi_i|W_i]$ would be nonparametrically identified and (5) would be a model with generated regressors entirely compatible with the framework of AIPR. Consider instead a situation in which $E_i[\xi_i|W_i]$ is not identified, either because we do not wish to impose rational expectations or because ξ_i is unobserved (assume that W_i is observable in this example). Suppose however that we assume

$$E[L_i|W_i] \leq E_i[\xi_i|W_i] \leq E[U_i|W_i] \quad \text{w.p.1,}$$

where a random sample of (L_i, U_i, W_i) is observable. This presupposes that, even though individual expectations are unknown and not identified, they are bounded w.p.1 by $E[L_i|W_i]$ and $E[U_i|W_i]$, perhaps owed to the fact that it is common knowledge among all individuals that $L_i \leq \xi_i \leq U_i$ w.p.1, and their expectations (however incorrect) are consistent with this property. This model is compatible with our general description with $\phi_i = F(E_i[\xi_i|W_i])$, $\underline{g}(W_i) = E[L_i|W_i]$ and $\overline{g}(W_i) = E[U_i|W_i]$.

A competitive entry model with possibly incorrect beliefs

Consider a strategic competition model between two competitors, labeled p, q (e.g, Walmart and Kmart). Suppose we are interested in some "outcome" or optimal decision by each one of these competitors (e.g, store size). For market i we label this outcome as Y^*_{pi}. Outcomes can be observed if and only if p decided to enter (i.e, compete) market i. Suppose the model is described as follows,

$$Y^*_{pi} = X_i'\beta^p_0 + \eta_{pi}, \quad d_{pi} = \mathbb{1}\left\{W_i'\delta^p_0 + \Delta(W_i, \lambda^p_0) \cdot E_{pi}[d_{qi}|W_i] + \varepsilon_{pi} \geq 0\right\}, \quad Y_{pi} = \begin{cases} Y^*_{pi} & \text{if } d_{pi} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

W_i here denotes the market's observable (to the econometrician) characteristics that determine entry decisions. In what follows we assume no exclusion restrictions between X_i and W_i. The expectation $E_{pi}[d_{qi}|W_i]$ denotes p's beliefs –unobserved in the data– for the probability that its competitor, q will decide to enter market i. The realization of W_i is assumed to inform p's beliefs, but the subscript E_{pi} in (6) implies that beliefs may still be random even conditional on W_i. However, assume in this example that beliefs are independent of[4] ε_{pi}. The natural setting for such an assumption is one where ε_{pi} is private information for p and independent of ε_{qi}. The term $\Delta(W_i, \lambda^p_0)$ represents a parametric function (indexed by parameter vector λ^p_0) that captures the strategic interaction effect.

Suppose we assume that $(\eta_{pi}, \varepsilon_{pi}) \perp (W_i, E_{pi}[d_{qi}|W_i])$, then we have a special case of a sample selection model, where $E\left[Y_{pi}|W_i, d_{pi} = 1\right] = X_i'\beta^p_0 + E\left[\eta_{pi}|\varepsilon_{pi} \geq -W_i'\delta^p_0 - \Delta(W_i, \lambda^p_0)E_{pi}[d_{qi}|W_i]\right]$. , Then for uncensored observations Y_i we can express

$$Y_{pi} = X_i'\beta^p_0 + \underbrace{m(W_i'\delta^p_0 + \Delta(W_i, \lambda^p_0) \cdot E_{pi}[d_{qi}|W_i])}_{=\phi_{pi}} + \nu_{pi}. \tag{7}$$

As we did in the example described in (4), suppose that the joint distribution of ε_{pi} and η_{pi} is assumed to be such that $m(\cdot)$ is *nonincreasing* (again, a special case would be bivariate Normality with positive correlation). Now suppose for simplicity that entry decisions are considered strategic substitutes[5], so $\Delta(W_i, \lambda^p_0) \leq 0$ w.p.1. Also suppose we parameterize the distribution of ε_p as $F(\cdot)$, and for simplicity assume this distribution is symmetric around zero (e.g, F is the Standard Normal distribution); the joint distribution of (η_p, ε_p) can remain nonparametric while maintaining the assumption that $m(\cdot)$ is nonincreasing.

In this example we assume beliefs are unobserved in the data. If we followed the vast majority of existing literature and assume rational expectations and Bayesian-Nash equilibrium (BNE) behavior (e.g, Seim (2006), Tamer (2003), Aradillas-Lopez (2010)), we could recover p's unobserved beliefs by solving the BNE conditions. Furthermore, if we assume that the BNE is either unique or that the equilibrium selection mechanism always chooses the same equilibrium we would have $E_{pi}[d_{qi}|W_i] = E[d_{qi}|W_i]$, and therefore beliefs could be estimated

[4] This is done for simplicity/brevity and can be relaxed.
[5] This can be easily relaxed to a setting where entry decisions are strategic substitutes in some markets and complements in others.

nonparametrically from the data. In this case, if we define

$$g(W_i;\gamma_0) = W_i'\delta_0^q + \Delta(W_i,\lambda_0^q)) \cdot E[d_{qi}|W_i],$$

then $g(W_i;\gamma_0) = g(W_j;\gamma_0)$ implies $\phi_i = \phi_j$ and pairwise matching methods like AIPR can be applied (as in Aradillas-Lopez (2012)). However, equilibrium behavior is a very strong assumption, as it presupposes that economic agents have perfect models about other agents. Following Aradillas-Lopez and Tamer (2008), suppose we replace BNE with the much weaker assumption that beliefs may be incorrect but they satisfy some basic notion of *rationality*. Note first that, regardless of how beliefs are constructed, we must have

$$0 \leq E_{pi}[d_{qi}|W_i] \leq 1.$$

By the strategic-substitutes assumption described above, this means that, regardless of how beliefs are constructed we must have

$$W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \leq W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot E_{pi}[d_{qi}|W_i] \leq W_i'\delta_0^p.$$

Suppose all we assume about p's beliefs is that they are consistent with this fact. Then we must have[6]

$$F(W_i'\delta_0^q + \Delta(W_i,\lambda_0^q)) \leq E_{pi}[d_{qi}|W_i] \leq F(W_i'\delta_0^q).$$

And therefore,

$$W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot F(W_i'\delta_0^q) \leq W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot E_{pi}[d_{qi}|W_i] \leq W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot F(W_i'\delta_0^q + \Delta(W_i,\lambda_0^q)). \quad (8)$$

Beliefs that satisfy these bounds are consistent with what Aradillas-Lopez and Tamer (2008) refer to as "Level-2 rationality". Recall from (7) that

$$\phi_i = m(W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot E_{pi}[d_{qi}|W_i]),$$

where $m(\cdot)$ is an unknown, nonincreasing transformation. If we relax the BNE restriction but assume Level-2 rationality, the resulting model is compatible with the general framework described in (2') with

$$\underline{g}_p(W_i;\gamma_0) = -\left(W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot F(W_i'\delta_0^q + \Delta(W_i,\lambda_0^q))\right),$$
$$\overline{g}_p(W_i;\gamma_0) = -\left(W_i'\delta_0^p + \Delta(W_i,\lambda_0^p) \cdot F(W_i'\delta_0^q)\right).$$

If the competitors in this model are at least "Level-2 rational" and the transformation m in (7) is nonincreasing, then the inequality in (2') is satisfied:

$$\underline{g}_p(W_i;\gamma_0) \geq \overline{g}_p(W_j;\gamma_0) \quad \Longrightarrow \quad \phi_{pi} \geq \phi_{pj}.$$

[6]Recall that we assume F to be symmetric around zero.

Level-k rationality, as defined in Aradillas-Lopez and Tamer (2008) is an iterative construction based on iterated deletion of dominated strategies. For example, if we go one step further and are willing to assume that competitors know that their opponents are Level-2 rational, and that their beliefs reflect this knowledge, we would now have

$$F\left(W_i'\delta_0^q + \Delta(W_i, \lambda_0^q) \cdot F(W_i'\delta_0^p)\right) \leq E_{pi}[d_{qi}|W_i] \leq F\left(W_i'\delta_0^q + \Delta(W_i, \lambda_0^q) \cdot F(W_i'\delta_0^p + \Delta(W_i, \lambda_0^p))\right).$$

These bounds correspond to Level-3 rationality in Aradillas-Lopez and Tamer (2008). From here the bounds $\underline{g}_p$ and $\overline{g}_p$ would now be

$$\begin{aligned}
\underline{g}_p(W_i;\gamma_0) &= -\left(W_i'\delta_0^p + \Delta(W_i, \lambda_0^p) \cdot F(W_i'\delta_0^q + \Delta(W_i, \lambda_0^q) \cdot F(W_i'\delta_0^p))\right), \\
\overline{g}_p(W_i;\gamma_0) &= -\left(W_i'\delta_0^p + \Delta(W_i, \lambda_0^p) \cdot F(W_i'\delta_0^q + \Delta(W_i, \lambda_0^q) \cdot F(W_i'\delta_0^p + \Delta(W_i, \lambda_0^p)))\right).
\end{aligned}$$

This construction has the advantage of always including BNE as a special case (i.e, BNE beliefs are always inside the Level-k bounds for any k) but it allows for inference that is robust to cases where competitors have incorrect beliefs. It also fits perfectly within the framework of examples described generically in (2').

3 Inference: an outline

Without the ability to match, the approach in AIPR (or those in the pairwise-differencing literature on control functions, such as Honoré and Powell (1994), Honoré and Powell (2005) or Aradillas-Lopez, Honoré, and Powell (2007)) cannot be applied. However, the model described in (2') produces moment inequalities on which inference can be based. Before we proceed let us denote $\theta \equiv (\beta, \gamma)$ as the full collection of parameters in our model, with γ_0 as the true value. While a number of moment-inequality inferential procedures are compatible with our setting, I will outline an approach that is explicitly based on the *conditional* nature of the moment inequalities. Take any iid pair (Y_1, X_1, W_1) and (Y_2, X_2, W_2) produced by a DGP as described in (2') and (3). Fix any (x, x') and define

$$\begin{aligned}
S(x,x';\theta) &= E\left[\left[(Y_1 - Y_2) - (X_1 - X_2)'\beta\right] \cdot \mathbb{1}\left\{\overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta)\right\} \,\middle|\, X_1 = x, X_2 = x'\right] \\
&= E\left[\left[(x - x')'(\beta_0 - \beta) + (\phi_1 - \phi_2) + (\nu_1 - \nu_2)\right] \cdot \mathbb{1}\left\{\overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta)\right\} \,\middle|\, X_1 = x, X_2 = x'\right].
\end{aligned}$$

Then,

$$\begin{aligned}
&S(x,x';\theta) \\
&= (x - x')'(\beta_0 - \beta) \cdot E\left[\mathbb{1}\left\{\overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta)\right\} \,\middle|\, X_1 = x, X_2 = x'\right] \\
&+ E\Big[\underbrace{E\left[(\phi_1 - \phi_2) \,\middle|\, \overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta), X_1 = x, X_2 = x'\right]}_{\leq 0} \cdot \mathbb{1}\left\{\overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta)\right\} \Big| X_1 = x, X_2 = x'\Big] \\
&+ E\Big[\underbrace{E\left[(\nu_1 - \nu_2) \,\middle|\, W_1, W_2, X_1 = x, X_2 = x'\right]}_{=0} \cdot \mathbb{1}\left\{\overline{g}(W_1;\gamma) \leq \underline{g}(W_2;\theta)\right\} \Big| X_1 = x, X_2 = x'\Big]
\end{aligned}$$

If the conditions in (2') and (3) are satisfied, we must have

$$S(x,x';\theta_0)\leq 0 \quad \text{for a.e } (x,x')\in Supp(X)^2.$$

In particular if we define

$$T(\theta)=E\left[\max\left\{S(X_1,X_2;\theta),0\right\}\right], \tag{9}$$

then we must have $T(\theta_0)=0$. Inference based on this restriction can be developed by extending the approach used in Aradillas-López and Gandhi (2016) and Aradillas-López, Gandhi, and Quint (2016), which look at models that yield functional inequalities of the type "$S(X;\beta_0)\leq 0$ w.p.1", which is equivalent to the mean-zero restriction "$E\left[\max\left\{S(X;\beta_0),0\right\}\right]=0$". Their methodology proposes constructing confidence sets (CS) for β_0 based on the functional $T(\beta)=E\left[\max\left\{S(X;\beta),0\right\}\right]$, for which they use estimators of the type

$$\widehat{T}(\beta)=\frac{1}{n}\sum_{i=1}^{n}\widehat{S}(X_i;\beta)\cdot\mathbb{1}\left\{\widehat{S}(X_i;\beta)\geq -b_n\right\},$$

where $b_n\longrightarrow 0$ is a sequence converging to zero at an appropriate rate. Under certain smoothness and regularity conditions, Aradillas-López and Gandhi (2016) and Aradillas-López, Gandhi, and Quint (2016) show that a test-statistic with pivotal properties can be constructed by properly normalizing $\widehat{T}(\beta)$, and that this procedure is both computationally simple to implement and it has good asymptotic properties, such as the ability to automatically adapt to the so-called *contact sets* (the regions where the conditional moment inequalities are binding). Analogously, in the case of (9) we can estimate

$$\widehat{T}(\theta)=\frac{1}{n\cdot(n-1)}\sum_{j\neq i}\sum_{i=1}^{n}\widehat{S}(X_i,X_j;\theta)\cdot\mathbb{1}\left\{\widehat{S}(X_i,X_j;\theta)\geq -b_n\right\}.$$

A CS for θ can be constructed by analyzing the asymptotic properties of $\widehat{T}(\theta)$ in a manner analogous to Aradillas-López and Gandhi (2016). This approach has the potential to be readily adapted to the case where the auxiliary parameter γ can be estimated from outside the model and plugged-in, in which case a CS for β can be based on

$$\widehat{T}(\beta)=\frac{1}{n\cdot(n-1)}\sum_{j\neq i}\sum_{i=1}^{n}\widehat{S}(X_i,X_j;\beta,\widehat{\gamma})\cdot\mathbb{1}\left\{\widehat{S}(X_i,X_j;\beta,\widehat{\gamma})\geq -b_n\right\}.$$

Under certain conditions, the conditional moment inequalities that result from (2') may be enough to point-identify θ (or β, in cases where the auxiliary parameter γ is identified outside the model). Estimation with conditional moment inequalities has been described, notably, in Khan and Tamer (2009). In this case a statistic $\widehat{T}(\theta)$ of the type described above can be employed to set-up the problem as an extremum estimator model.

4 Concluding remarks

The approach in AIPR –and in the pairwise-differencing literature as a whole– presupposes the ability to match (asymptotically) the control functions across pairs of observations. In this note I have argued that it is easy to construct examples where matching is not possible, either because of interval or missing data, or because of so-called incomplete models such as those arising from multiple equilibria (or some other solution concept such as rationalizability). In the examples studied here, even though matching is no longer feasible, the model still produces testable implications, in the form of conditional moment inequalities. Using recent advances in moment inequality models, inference can still be possible. This shows that the concept of control functions –an idea advocated and developed in James L. Powell's body of work– is a powerful one, capable of extending to settings with partial identification.

References

Ahn, H., H. Ichimura, J. Powell, and P. Ruud (2015). Simple estimators for intertible index models. Forthcoming, JBES.

Ahn, H. and J. Powell (1993). Semiparametric estimation of censored selection models. *Journal of Econometrics 58*, 3–29.

Aradillas-Lopez, A. (2010). Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics 157*(2), 409–431.

Aradillas-Lopez, A. (2012). Pairwise difference estimation of incomplete information games. *Journal of Econometrics 168*(1), 120–140.

Aradillas-López, A. and A. Gandhi (2016). Estimation of games with ordered actions: An application to chain-store entry. *Quantitative Economics*. Forthcoming.

Aradillas-López, A., A. Gandhi, and D. Quint (2016). A simple test for moment inequality models with an application to english auctions. *Journal of Econometrics 194*(1), 96–115.

Aradillas-Lopez, A., B. Honoré, and J. Powell (2007). Pairwise difference estimation with nonparametric control variables. *International Economic Review 48*(4), 1119–1158.

Aradillas-Lopez, A. and E. Tamer (2008). The identification power of equilibrium in games. *Journal of Business and Economic Statistics 26*(3), 261–283.

Blundell, R. and J. Powell (2004). Endogeneity in semiparametric binary response models. *Review of Economic Studies 71*, 655–679.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153 – 161.

Hong, H. and M. Shum (2009). Pairwise difference estimator of a dynamic optimization model. *Review of Economic Studies, forthcoming*.

Honoré, B. and J. Powell (1994). Pairwise difference estimators of censored and truncated regression models. *Journal of Econometrics 64*(1-2), 241–278.

Honoré, B. and J. Powell (2005). Pairwise difference estimation in nonlinear models. In D. Andrews and J. Stock (Eds.), *Identification and Inference in Econometric Models. Essays in Honor of Thomas Rothenberg*, pp. 520–553. Cambridge University Press.

Imbens, G. and W. Newey (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica 77*, 1481–1512.

Khan, S. and E. Tamer (2009). Inference on endogenously censored regression models using conditional moment inequalities. *Journal of Econometrics 152*(2), 104–119.

Manski, C. and E. Tamer (2002). Inference on regressions with interval data on a regressor or outcome. *Econometrica 70*, 519–547.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics 37*(3).

Tamer, E. T. (2003). Incomplete bivariate discrete response model with multiple equilibria. *Review of Economic Studies 70*, 147–167.