Testing Multivariate Economic Restrictions using Quantiles: The Example of Slutsky Negative Semidefiniteness

Stefan Hoderlein, Boston College

(Joint with Holger Dette and Natalie Neumayer)

- **Objective of the Paper**: 1. Testing shape constraint imposed by rationality in heterogeneous population with complex unobservable.

- 2. Using distributional information/quantiles in systems of equations without triangularity/monotonicity.

Difficulties with this Objective

- Problem 1: Relationship between quantiles and structural unobservables in the absence of scalar monotonicity.

- Problem 2: How to use an univariate concept (quantile) in systems of equations.

- Problem 3: Interaction between problems 1 and 2. What added complications arise if we consider structural unobservables in systems of equations.

Literature

- Shape constraints and heterogeneity: Matzkin (1994, 2007)

- Testing rationality + parametric: Stone (1954), Deaton and Muellbauer (1980), Blundell et al. (1994). Control for observed heterogeneity

- Integrability constraints and nonparametric mean regressions: Stoker (1989), Hildenbrand et al (1991), Lewbel (2001).

- Nonparametric testing literature. We could do omission of variables.

Literature

- Nonseparable models, Chesher (2003), Altonji and Matzkin (2005), Imbens and Newey (2009).

- Particularly close: Hoderlein and Mammen (2007, 2009)

- Mean Regressions and testing: Hoderlein (2005, 2010).

- Alternative: Weak Axiom, Blundell, Kristensen and Matzkin (2010), Hoderlein and Stoye (2010). Problems of triangular structure.

Overview of today's talk

- 1. Motivation/Literature
- 2. Model
- 3. Identification Result
- 4. Test statistic and Bootstrap.
- 5. Simulation/Application

(Unobserved) Population Model

Assumption 1, DGP: "Nonseparable" System of Demand Equations

$$\begin{aligned} Y_1 &= \phi_1(P, X, U) \\ Y_2 &= \phi_2(P, X, U) \\ &\vdots \\ Y_L &= \phi_L(P, X, U), \end{aligned}$$

$Y_l =$ demand for good l, $Y \in \mathcal{Y} \subseteq \mathbb{R}_+^L$

$P = (P_1, ..., P_L)' \in \mathcal{P} \subseteq \mathbb{R}_+^L$, L- vector of prices.

$X \in \mathcal{X} \subseteq \mathbb{R}_+$, "income",

$U \in \mathcal{U} \subseteq \mathbb{R}^\infty$, "preferences", high (potentially countably infinite) dimensional unobservable.

(Unobserved) Population Model

- What do we allow for: all equations may depend on $d > L$ unobservables in arbitrary fashion.

- We do not assume triangular structure.

- We do not assume that individuals are of the same type.

- There could be many types, with many parameters each.

(Unobserved) Population Model

- "Treatment effects" scenario - "excess" heterogeneity, ϕ not identifiable.

- However: Averages are identified.

- **Economic theory**: Under standard assumptions on utility, ϕ smooth, differentiable in p, x.

- For every $U = u$, Slutsky matrix negative semidefinite.

Slutsky matrix.

$$\mathfrak{S}(P,X,U) = D_p\phi(P,X,U) + \partial_x\phi(P,X,U)\phi(P,X,U)',$$

- D_p matrix of partial derivatives, ∂_x vector. Has to be nsd, i.e.,

$$\mathfrak{S}(P,X,U) \leq 0, \qquad \mathbb{P}_{PXU} - a.s.$$

- Key restriction of utility maximization. Equivalent to Weak Axiom.

Independence Assumption

- Problem: in heterogeneous population, depends on (unobserved) U.

- Need to make statement about how this relates to observables

- Independence condition (no covariates). Exogenous case

$$F_{U|PX} = F_U$$

Independence Assumption

- With covariates/household characteristics $S \in \mathcal{S} \subseteq \mathbb{R}^l$. Let $U = \vartheta(S, A)$. $A \in \mathcal{A} \subseteq \mathbb{R}^\infty$.

- Independence condition. Exogenous case

$$F_{U|PXS} = F_{U|S}$$

- "Selection on observables".

Independence Assumption

- With Endogeneity of (w.l.o.g. X): Control function approach. Assume there exist instruments Z such that

$$X = \mu(P, S, Z) + V,$$

(or scalar monotonicity, see Chesher (2003), Hoderlein and Vanhems (2010)), and

$$F_{U|PXSV} = F_{U|SV}$$

- **Assumption 2**: "Selection on observables and controls".

Other Ingredients

- Regularity conditions on ϕ : smoothness, uniform bounds on derivatives (“**Assumption 3**”)

- Innovation: use distributional information to learn about $\mathfrak{S}(P, X, U)$.

- Use regression quantiles, C random scalar, D random vector. Formally define $k^{\alpha}_{C|D}(d)$ by

$$\mathbb{P}(C \leq k_{\alpha}(d)|D = d) = \alpha,$$

- Regularity conditions on certain quantiles (“**Assumption 4**”)

Other Ingredients

- Straightforward corollary from Hoderlein and Mammen (2007)

- *Let A1 - A4 hold, then for any* Y_l, $l = 1, .., L$,

$$\begin{aligned} \partial_x k^{\alpha}_{Y_l|PXSV}(p,x,s,v) &= \\ \mathbb{E}[\partial_x \phi(P,X,U)|P &= p, X = x, S = s, V = v, Y = k^{\alpha}_{Y_l|PXSV}(x)], \end{aligned}$$

and analogously for $D_p\phi$.

Other Ingredients

- Corollary - Interpretation

  $$\partial_x k^{\alpha}_{Y_l|PXSV}(P,X,S,V) = \mathbb{E}[\partial_x \phi(P,X,U)|P,X,S,V,Y = k^{\alpha}_{Y_l|PXSV}].$$

- Best approximation to unobserved effect of interest $\partial_x \phi$, given observed information.

- Effect of a small treatment, for subpopulation defined by treatment intensity $P = p, X = x$ and proxies for preferences $S = s, V = v, Y = k^{\alpha}_{Y_l|PXSV}(p,x,s,v)$.

Problem

- Univariate result in a "deep" sense:

- Example, two equations

$$
\begin{aligned}
Y_1 &= \phi_1(X, A), \\
Y_2 &= \phi_2(X, A)
\end{aligned}
$$

- Hoderlein and Mammen (2009): $\mathbb{E}[\partial_x \phi_1(X, U)|X, Y_1, Y_2]$ not (point) identified from distribution of data.

- No hope for direct "systems of equations" attack using all L quantiles.

Main identification result in this paper

- Some preliminaries: Given $L-1$ demands, the L-th is determined by budget constraint: omit equation L

- We impose homogeneity of degree zero (one price drops out, all prices are relative)

- Hence also only $L-1$ prices.

Main identification result

- Key identification idea in DHN: reduce multivariate problem to set of univariate problems, akin to the Cramer-Wold device.

- In particular, form indices of dependent variable.

- Formally, let $Y(b) = b'Y = b'\phi$, for

- Let $k(\alpha, b|w)$ denote the conditional α quantile of $Y(b)$ given $W = (P, X, S, V)$

- Let ∇_x denote the gradient.

Main identification result

Theorem 1: *Let assumptions A1–A4 hold. Then,*

$$\mathfrak{S}(p,x,u) \;\; nsd \Rightarrow \nabla_p k(\alpha,b|w)'b + \partial_x k(\alpha,b|w)k(\alpha,b|w) \leq 0$$

for all $(\alpha,b) \in (0,1) \times \mathbb{S}_{L-1}$, *and all* $(w,u) \in \mathcal{W} \times \mathcal{U}$.

- **Main identification result**

- **Remarks:** No triangular structure, no monotonicity, still learn something by looking at implications on set of distributions.

- $g(\phi(\cdot,\cdot)) \in \mathcal{B} \Rightarrow g^*(F_{Y|PX}(\cdot,\cdot)) \in \mathcal{B}'$.

- Logic $\neg\ g^*(F_{Y|PX}(y,p,x)) \Rightarrow \neg\ g(\phi(y,p,u))$ with $\mathbb{P}\left[Y = y, P = p, U = u\right] > 0$.

- We may learn something: For which value of $(p, x; \alpha, b)$, i.e., subpopulation, rationality fails. Important for refining economic theory. Example with covariates. $S =$ urban two person households.

Main identification result

- **Remarks:**

- Cramer Wold device: (conditional distribution of all linear combinations used $\rightarrow$ use entire conditional distribution of the data).

- But identifies only set of conditional expectations with σ-algebras: $\mathbb{E}[\cdot|P, X, S, V, Y(b)]$ indexed by b.

- **Proof** uses main idea of HM, combined with directional derivatives

Test Statistic

- Introduce

$$k_{p_\ell}(\alpha, b \mid w) = \partial_{p_\ell} k(\alpha, b \mid w)\Big|_{(p,x)=w}$$
$$k_x(\alpha, b \mid w) = \partial_x k(\alpha, b \mid w)\Big|_{(p,x)=w}$$

- H_0 : $\quad \forall \alpha \in A, b = (b_1, \ldots, b_{L-1})' \in \mathbb{S}_{L-1}$:

$$\sum_{\ell=1}^{L-1} b_\ell k_{p_\ell}(\alpha, b \mid w) + k_x(\alpha, b \mid w) k(\alpha, b \mid w) \leq 0$$

Test Statistic

- Define

$$T_n = \sqrt{nh^{L+2}} \sup_{\alpha \in A, b \in \mathbb{S}_{L-1}} R_n(\alpha, b \mid w), \tag{1}$$

where

$$R_n(\alpha, b \mid w) = \sum_{\ell=1}^{L-1} b_\ell \widehat{k}_{p_\ell}(\alpha, b \mid w) + \widehat{k}_x(\alpha, b \mid w)\widehat{k}(\alpha, b \mid w). \tag{2}$$

Test Statistic

- Let $\tau_\alpha(u) = u(\alpha - I\{u < 0\})$

- Estimators are local polynomial quantile estimators:

$$
\begin{aligned}
& (\widehat{\mu}_0, \widehat{\mu}_1, \widehat{\mu}_2) \\
= \; & \arg \min_{(\mu_0,\mu_1,\mu_2)\in\mathbb{R}\times\mathbb{R}^L\times\mathbb{R}^{L\times L}} \sum_{i=1}^{n} \tau_\alpha\Big(Y_i'b - \mu_0 - \mu_1'(W_i - w) \\
& -(W_i - w)'\mu_2(W_i - w)\Big) K\Big(\frac{W_i - w}{h}\Big)
\end{aligned}
$$

- Then, $\widehat{k}(\alpha, b \mid w) = \widehat{\mu}_0$, $\widehat{k}_{p_\ell}(\alpha, b \mid w) = \widehat{\mu}_{1,\ell}$, $\ell = 1, \ldots, L-1$, and $\widehat{k}_x(\alpha, b \mid w) = \widehat{\mu}_{1,L}$.

Test Statistic

- Under standard assumptions, $R_n(\alpha, b \mid w)$ consistently estimates

$$R(\alpha, b \mid w) = \sum_{\ell=1}^{L-1} b_\ell k_{p_\ell}(\alpha, b \mid w) + k_x(\alpha, b \mid w) k(\alpha, b \mid w)$$

and we have the following weak convergence result.

Test Statistic

- **Theorem 2:** *Let $f_{P,X}$ denote the density of (P,X), and let $f_{Y(b)|P,X}(\cdot \mid w)$, $F_{Y(b)|P,X}(\cdot \mid w)$ denote the conditional density and distribution function of $Y(b)$, given $(P,X) = w$. Under assumptions specified in the paper, the process*

$$\sqrt{nh^{L+2}}\Big(R_n(\alpha, b \mid w) - R(\alpha, b \mid w)\Big)_{\alpha \in A, b \in \mathbb{S}_{L-1}}$$

converges (for w fixed) weakly to a Gaussian process $G(\alpha, b \mid w)_{\alpha \in A, b \in \mathbb{S}_{L-1}}$ with covariance

Test Statistic

- **Theorem 2 (cont.):** *Gaussian process* $G(\alpha, b \mid w)_{\alpha \in A, b \in \mathbb{S}_{L-1}}$ *with covariance*

$$\mathrm{Cov}\Big(G(\alpha, b \mid w), G(\tilde{\alpha}, \tilde{b} \mid w)\Big)$$
$$= \Big[\mathbb{P}\Big(Y(b) \leq k(\alpha, b \mid w), Y(\tilde{b}) \leq k(\tilde{\alpha}, \tilde{b} \mid w) \,\Big|\, (P, X) = w\Big)$$
$$- F_{Y(b)|P,X}\Big(k(\alpha, b \mid w) \,\Big|\, w\Big) F_{Y(\tilde{b})|P,X}\Big(k(\tilde{\alpha}, \tilde{b} \mid w) \,\Big|\, w\Big)\Big]$$
$$\times \frac{\int K^2(\overline{p}, \overline{x})(b'\overline{p} + k(\alpha, b \mid w)\overline{x})(\tilde{b}'\overline{p} + k(\tilde{\alpha}, \tilde{b} \mid w)\overline{x})\, d(\overline{p}, \overline{x})}{f_{Y(b)|P,X}(k(\alpha, b \mid w) \mid w) f_{Y(\tilde{b})|P,X}(k(\tilde{\alpha}, \tilde{b} \mid w) \mid w) f_{P,X}(w)(\int u^2 \kappa}$$

.

Test Statistic

- Only get distributional result for centered statistic

$$\tilde{T}_n = \sqrt{nh^{L+2}} \sup_{\alpha,b} (R_n(\alpha, b \mid w) - R(\alpha, b \mid w)) \tag{3}$$

both under the null hypothesis and under fixed alternatives.

- Corollary to theorem 2: For each $c \in \mathbb{R}$,

$$\mathbb{P}(\tilde{T}_n > c) \overset{n\to\infty}{\longrightarrow} \mathbb{P}(\sup_{\alpha,b} G(\alpha, b) > c). \tag{4}$$

Test Statistic

- Structure of the null hypothesis, $H_0 : R(\cdot, \cdot \mid w) \leq 0$

- It follows that $\mathbb{P}(T_n > c) \leq \mathbb{P}(\tilde{T}_n > c)$ under H_0,

- Hence, obtain asymptotically level γ test by rejecting H_0 whenever $T_n > c_\gamma$, where $\mathbb{P}(\sup_{\alpha,b} G(\alpha, b) > c_\gamma) = \gamma$.

Test Statistic

- Pointwise tests (ie, conditional on fixed values of regressors) as in Hoderlein (2005, 2010).

- Objective: Get fraction of population for which rationality is rejected through Honore-Mueller correction.

$$\mathbb{P}_X\left[reject\right] = \mathbb{P}_X\left[not\ rational\right] + 0.05\left[1 - \mathbb{P}_X\left[not\ rational\right]\right]$$

- Solve for $\mathbb{P}_X\left[not\ rational\right]$.

Test Statistic

- If interest is in joint null hypothesis of rationality at m specified values of x.

- Consider the largest t-statistic test statistic. Adjust critical value upwards.

- Specifically, if the individual t-tests are of level α, the level α^* of the overall test satisfies $(1-\alpha)^m = 1-\alpha^*$

Test Statistic - Bootstrap

- Asymptotic distribution to involved, hence get bootstrap, $k = 1, .., m$.

- Keep covariates and define $(P_k^*, X_k^*) = (P_k, X_k)$.

- For each fixed covariate (P_k, X_k) we generate Y_k^* from conditional distribution of Y, given $(P, X) = (P_k, X_k)$, i. e. $F_{Y|P,X}(\cdot \mid P_k, X_k)$.

- For one-dimensional Y this method coincides with the bootstrap procedure suggested in HM (2009).

Simulation

- DGP Let $A \sim \mathcal{U}[0,1]$. Then, simple model with no income effect:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 0.9A \\ 0.5A \\ 0.7A \end{bmatrix} + \underbrace{\begin{bmatrix} -0.25A + \lambda & 0.1A & 0.1A \\ 0.1A & -0.25A & 0.1A \\ 0.1A & 0.1A & -0.25A \end{bmatrix}}_{Slutsky} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix},$$

- For $\lambda = 0$ always nd for entire population.

Simulation

- DGP Let $A \sim \mathcal{U}[0,1]$. Then, model such that Slutsky matrix:

$$S(A) = \begin{bmatrix} -0.25A + \lambda & 0.1A & 0.1A \\ 0.1A & -0.25A & 0.1A \\ 0.1A & 0.1A & -0.25A \end{bmatrix},$$

- As λ increases, parts of population not nd. For $\lambda = 0.05$, $5/25 = 0.2$ of population not rational.

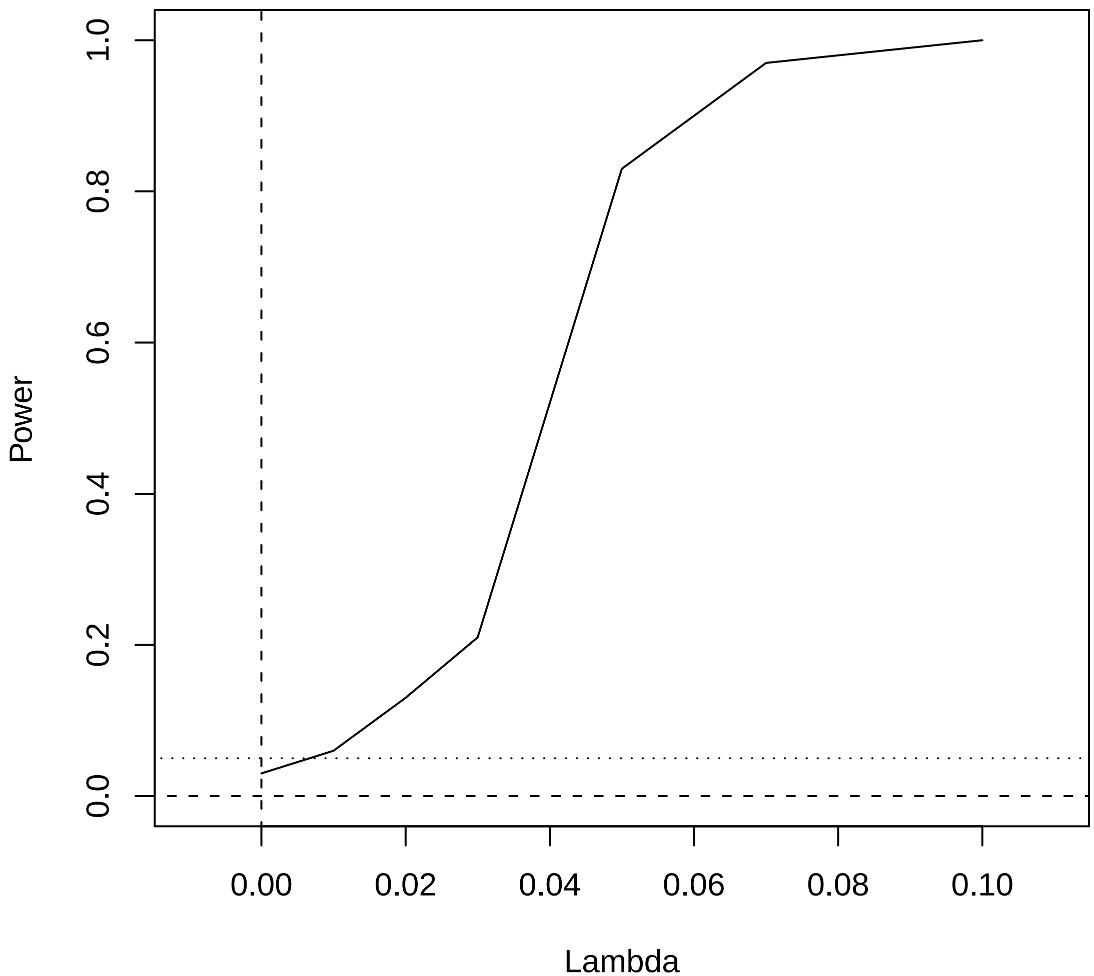
Simulation

- Numbers, $n = 2000$ observations

- Data simulated from normal distributions, but same means and variances as data in application.

- Results. Size = 0.03 (expected size distortion due to non-centering).

λ	0.01	0.02	0.03	0.04	0.05	0.07	0.1
% *Population not Rational*	0.04	0.08	0.12	0.16	0.20	0.28	0.40
Power	0.08	0.15	0.30	0.62	0.83	0.97	0.999


1.0
0.8
0.6
0.4
0.2
0.0
Power
0.00
0.02
0.04
0.06
0.08
0.10
Lambda

Simulation

- Similar, but even stronger results when difference between rational and not rational bigger.

- Models with income effect: similar results.

- To do: more alternatives etc.

Applications

- British FES data (repeated cross sections, thanks to Richard Blundell)

- Repeated cross sections.

- Construct Stone-Lewbel (cross section) prices as in Lewbel (1995), see also Hoderlein and Mihaleva (2008).

- Preliminary results: Find rationality largely not rejected,

$$\mathbb{P}_X\left[Rational\right] = 0.92.$$

  No obvious structure in violations.

Summary/Outlook

- Showed how do conduct hypothesis testing in a scenario with:

- Complicated (“excess”) Heterogeneity

- Systems of Equations

- Using entire distribution of the data (b-indexed quantiles, Cramer-Wold device)

Summary/Outlook

- Proposed test statistic.
- Established large sample theory
- Derived bootstrap method.
- Worked well in simulations: able to detect even relatively small fractions of almost rational individuals
- Found in application support for core of utility maximization.

Summary/Outlook

- Principle can be applied to omission of variables
- Also in systems of simultaneous equations/Panels.
- Testing for endogeneity in a control function fashion.
- We focus on most important economic shape constraint.