Identification of Nonseparable Triangular Models with Discrete Instruments: Supplementary Materials

Xavier D'Haultfœuille *   Philippe Février†

December 2014

Abstract

In these supplementary materials, we first discuss the link with group theory and the freeness and nonfreeness properties. We then discuss the extension to a multivariate X. The third section gathers all proofs.

1 Link with group theory

1.1 Definitions

We first recall some useful definitions on group theory. A group $\mathcal{S}$ is a set endowed with a binary operator $*$ which satisfies three properties. The first is associativity: for all $(s_1, s_2, s_3) \in \mathcal{S}^3$, $(s_1 * s_2) * s_3 = s_1 * (s_2 * s_3)$. The second is the existence of an identity element $e \in \mathcal{S}$ satisfying $s * e = e * s = s$ for all $s \in \mathcal{S}$. The third is the existence of inverses. Every element $s \in \mathcal{S}$ admits an element called its inverse and denoted s^{-1} which satisfies $s * s^{-1} = s^{-1} * s = e$.

*CREST. E-mail address: xavier.dhaultfoeuille@ensae.fr.
†CREST. E-mail address: fevrier@ensae.fr.

The set $\mathcal{B}$ of all bijections onto $\mathcal{X}$, endowed with the composition operator, is an example of a group.

A group $\mathcal{S}$ is said to be abelian if for every $(s_1, s_2) \in \mathcal{S}^2$, $s_1 * s_2 = s_2 * s_1$. A subgroup $\mathcal{T}$ of $\mathcal{S}$ is a subset of $\mathcal{S}$ which is itself a group for $*$. If we let $(\mathcal{T}_i)_{i\in\mathcal{I}}$ denote a family of subgroups of $\mathcal{S}$, one can check that $\cap_{i\in\mathcal{I}}\mathcal{T}_i$ is also a group. The group generated by a subset I of $\mathcal{S}$ is the intersection of all subgroups of $\mathcal{S}$ containing I. By definition, it is the smallest subgroup of $\mathcal{S}$ including I. In the paper, $\mathcal{S}$ is the subgroup of $\mathcal{B}$ generated by the functions $(s_{ij})_{(i,j)\in\{1,\dots,K\}^2}$.

We also define the notion of group actions and orbits. For any set $\mathcal{A}$ and a group $\mathcal{S}$, a group action . is a function from $\mathcal{S} \times \mathcal{A}$ to $\mathcal{A}$ (denoted by $s.x$) satisfying, for every $(s_1, t) \in \mathcal{S}^2$ and $x \in \mathcal{A}$, $(s_1 * t).x = s_1.(t.x)$ and $e.x = x$. The orbit $\mathcal{O}_x$ of $x \in \mathcal{A}$ is then defined by

$$\mathcal{O}_x = \{s.x, s \in \mathcal{S}\}.$$

In the paper, the group action is $s.x = s(x)$ and the orbit of x is the set $\mathcal{O}_x = \{s(x), s \in \mathcal{S}\}$. Finally, a group action . is free if $s.x = x$ for some $x \in \mathcal{A}$ implies that $s = e$. This definition coincides, in the setting of the paper, with the freeness property.

1.2 The freeness and nonfreeness properties

1.2.1 General results

Let us recall that the freeness properties holds if there exists no $s \in \mathcal{S}$ different from the identity function that admits a fixed point. The nonfreeness property holds if there exists $s \in \mathcal{S}$ different from the identity function that admits a positive and finite number of fixed points.

Whether these properties hold depend on the way the instrument Z affects the endogenous variable X. A first observation is that if there exists $(i, j) \in \{1, ..., K\}^2$ such that $F_{X|Z=i}$ and $F_{X|Z=j}$ cross at least once and at most a finite number of times on $\mathcal{X}$, then the nonfreeness holds.[1] A case where the different $(F_{X|Z=i})_{i\in\{1,\dots,K\}}$ cross are generalized location scale-models, with the exception of pure location models. In the latter case, the freeness property holds.

Proposition S1 *Suppose that*

$$h(Z,\eta) = \mu(\nu(Z) + \sigma(Z)\eta), \qquad (1.1)$$

where $Z \perp\!\!\!\perp \eta$, *Support*$(\eta) = \mathbb{R}$, $\sigma(Z) > 0$ *and* μ *is a strictly increasing function from* $\mathbb{R}$ *to* $\mathcal{X}$. *If* $\sigma(Z)$ *is not constant, the nonfreeness property holds. Otherwise, the freeness property holds.*

1.2.2 Illustration

Let us illustrate the freeness and nonfreeness properties in a specific context. Suppose that we are interested in measuring the effect of unemployment duration X on an health index Y, using policy changes on unemployment benefits as an instrument Z. Suppose that the hazard rate of X conditional on $Z = z$ satisfies a Cox model $\lambda_z(t) = \lambda_0 \exp(-cb_z(t))$, where $b_z(t)$ denotes unemployment benefits at date t under policy status z.[2] We show hereafter that depending on the type of policy changes that we consider, we end up with either freeness or nonfreeness.

First , if the unemployment benefits are less generous after the policy change, so that $b_1(t) < b_2(t)$ for all t, $F_{X|Z=2}$ stochastically dominates $F_{X|Z=1}$. The freeness property holds because all unemployed people have less incentives to

[1]Torgovitsky (2014) uses such fixed points to achieve identification. Interestingly, this crossing property is also used by Guerre et al. (2009) to achieve identification of a first-price auctions models with risk averse bidders. They use for that purpose exogenous variation in the number of bidders, which plays the role of discrete instrument in their framework. The crossing they use is on the bidding functions, and is automatically satisfied by the theoretical bidding model.

[2]This toy model is useful to discuss the economic contents of our assumptions but does not pretend to be fully realistic.

find a job. Now, consider the case where unemployment benefits were initially constant over time, $b_1(t) = b_1$, but then become decreasing: $b_2(t) = b_{21}\mathbb{1}\{t \leq t_0\} + b_{22}\mathbb{1}\{t > t_0\}$ for a given threshold t_0, with $b_{21} > b_1 > b_{22}$. The new policy is thus more generous for short periods of unemployment, and less generous for longer ones. The integrated hazard satisfies $\Lambda_1(t) = \int_0^t \lambda_1(u)du = B_1 t$ and $\Lambda_2(t) = B_{21}t\mathbb{1}\{t \leq t_0\} + (B_{22}t + (B_{21} - B_{22})t_0)\mathbb{1}\{t > t_0\}$, with $B_1 = \lambda \exp(-cb_1)$ and $B_{2i} = \lambda \exp(-cb_{2i})$, $i \in \{1, 2\}$. Because $F_{X|Z=z}(x) = 1 - \exp(-\Lambda_z(x))$ and $s_{12} = F^{-1}_{X|Z=2} \circ F_{X|Z=1}$, we obtain $s_{12} = \Lambda_2^{-1} \circ \Lambda_1$. Hence, nonfreeness holds here, because Λ_1 and Λ_2 cross once.

Now, suppose that we experience several changes, but $b_z(t) = a_z + b(t)$, so that unemployment benefits differ by the same constant over time under the different policies. Then X satisfies a generalized location model, so that the freeness property holds by Proposition 1. Finally, we provide an example where $K \geq 3$ and nonfreeness holds, though the $(s_{ij})_{i,j}$ do not cross. Suppose that $b_1(.)$ and $b_2(.)$ are as previously but $b_{21} < b_1$, so that the second policy is always less generous than the first. Suppose also we have a third policy satisfying $b_3(t) = b_{31}\mathbb{1}\{t \leq t_0\} + b_{22}\mathbb{1}\{t > t_0\}$ and $b_{31} < b_{21}$. The third policy is thus less generous than the second one. As a result, $s_{ij}(x) > x$ for all $i < j$. Suppose also for the ease of exposition that $b_{21} = 2b_1/3$, $b_{31} = b_1/4$ and $b_{22} = b_1/5$, and let us prove that $s_{31} \circ s_{12}^2$ admits a unique fixed point.[3]

Within this framework, the integrated hazard rates satisfy $\Lambda_1(t) = B_1 t$, with $B_1 = \lambda \exp(-cb_1)$, and

$$\Lambda_j(t) = B_{j1}t\mathbb{1}\{t \leq t_0\} + (B_{22}t + (B_{j1} - B_{22})t_0)\mathbb{1}\{t > t_0\}$$

for $j = 2, 3$, with $B_{ji} = \lambda \exp(-cb_{ji})$. Because $s_{ij} = \Lambda_j^{-1} \circ \Lambda_i$, it follows that

$$s_{i1}(t) = E_{i1}t\mathbb{1}\{t \leq t_0\} + (Et + (E_{i1} - E)t_0)\mathbb{1}\{t > t_0\}$$

[3]The same result holds for more general values of (b_{21}, b_{31}, b_{22}) but the argument is more complicated.

for $i = 2, 3$, with $E_{i1} = B_{i1}/B_1 > 1$ and $E = B_{22}/B_1 > E_{i1}$. Some computations yield

$$\begin{aligned} s_{21}^2(t) &= E_{21}^2 t \mathbb{1}\{t \leq t_0/E_{21}\} + (E_{21}Et + (E_{21} - E)t_0)\mathbb{1}\{t_0 \geq t > t_0/E_{21}\} \\ &\quad + (E^2 t + (E_{21} - E)(E+1)t_0)\mathbb{1}\{t > t_0\}. \end{aligned}$$

We have $E_{31} = \exp(3cb_1/4) > \exp(2cb_1/3) = E_{21}^2$. Thus $s_{31}(t) > s_{21}^2(t)$ for $t \in (0, t_0/E_{21})$ and they do not cross on this interval. The functions s_{31} and s_{21}^2 are then linear on $(t_0/E_{21},\ t_0)$ and $s_{31}(t_0/E_{21}) > s_{21}^2(t_0/E_{21})$. If $s_{31}(t_0) > s_{21}^2(t_0)$, the functions do not cross on this interval either. Otherwise, if $s_{31}(t_0) < s_{21}^2(t_0)$, the functions cross only once on the interval $(t_0/E_{21},\ t_0)$. Finally, s_{31} and s_{21}^2 are linear on $[t_0, +\infty)$ with different slopes. Moreover, for $t > t_0$, $E^2 > E$ so that $s_{21}^2(t) > s_{31}(t)$ for $t > t_0$ large enough. Hence, if $s_{31}(t_0) > s_{21}^2(t_0)$, the functions cross only once on the interval $[t_0, +\infty)$ whereas they do not cross if $s_{31}(t_0) < s_{21}^2(t_0)$. At the end, in all cases, $s_{31} \circ s_{12}^2$ admits a unique fixed point. The nonfreeness property holds though the $(s_{ij})_{i,j}$ do not cross.

2 The multivariate case

In the multivariate case, the topology of the orbits is more complicated and a full classification is difficult to obtain. Yet, Theorem 1 is still valid and previous ideas can be partially extended. We first write the suitable generalizations of Assumptions 2-3 in this context. Henceforth, $\mathcal{H}$ still denotes the interior of the support of η.

Assumption S1 *(Dual strict monotonicity) $\varepsilon \in \mathbb{R}$, $h(Z, \eta) = (h_1(Z, \eta_1), ..., h_d(Z, \eta_d))$ and for all $(x, z, m) \in \mathcal{X} \times \{1, ..., K\} \times \{1, ...d\}$, $\tau \mapsto g(x, \tau)$ and $v \mapsto h_m(z, v)$ are strictly increasing.*

Assumption S2 *(technical restrictions)*
(i) Support$(X|Z = z) = \prod_{m=1}^d [\underline{x}_m, \overline{x}_m]$ with $-\infty \leq \underline{x}_m < \overline{x}_m \leq \infty$ independent of z.

(ii) ε has a uniform distribution.
(iii) F_{η_m} is continuous and strictly increasing on its support for $m \in \{1, ..., d\}$.
(iv) $(u, v) \mapsto F_{\varepsilon|\eta=v}(u)$ is continuous and strictly increasing in u for all $v \in \mathcal{H}$.
(v) $g(.,.)$ and $h(z,.)$ are continuous on $\mathcal{X} \times (0,1)$ and $\mathcal{H}$ respectively.

It is easy to see that under these conditions, Theorem 1 is still valid, where $\mathcal{S}$ is still the group generated by the $s_{ij}(x) = (s_{ij1}(x_1), ..., s_{ijd}(x_d))$, with $s_{ijm} = F^{-1}_{X_m|Z=j} \circ F_{X_m|Z=i}$. The issue is therefore whether the condition on the orbits holds or not.

2.1 The free case

The powerful tools that we used for the univariate free case, namely Hölder's and Denjoy's theorems, do not apply anymore. Hölder's theorem states that if freeness holds for a group of functions on the real line, this group is abelian. Thanks to this property, we can reduce our study to the unit circle. This result does not hold however for functions of several variables. Moreover, even if we were able to come back to the unit circle on each coordinate, Denjoy's theorem would only prove density on each of these coordinates but not on the cartesian product of these unit circles, which would be necessary to establish full identification.

Even if we cannot use the same proof as in the univariate case, the generalized location model still provides some interesting insights. Suppose that

$$X_m = \mu_m\left(\nu_m(Z) + \eta_m\right), \ m = 1...d \tag{2.1}$$

where μ_m is strictly increasing and continuous and, without loss of generality, $\nu_1(1) = ... = \nu_d(1) = 0$. We let hereafter A denote the $K-1 \times d$-matrix of typical $(k-1, m)$ element $\nu_m(k)$, for $k = 2, ..., K$, and A_k be the kth line of A. We make the following assumption.

Assumption S3 *(Rank and non-periodicity condition) (i) The matrix A has rank d and (ii) supposing without loss of generality that $(A_1, ..., A_d)$ are lin-*

early independent, there exists $i > d$ such that $A_i = \sum_{k=1}^{d} \lambda_k A_k$ and for all $(c_1, ..., c_d) \in \mathbb{Z}^d$, $(c_1, ..., c_d) \neq (0, ..., 0)$, $\sum_{k=1}^{d} \lambda_k c_k \notin \mathbb{Z}$.

Condition (i) is similar to the standard rank condition in linear IV models, and actually identical when $\mu_1, .., \mu_d$ are the identity function. Condition (ii) is similar to the non-periodicity condition imposed in Assumption 4 in the univariate case, and can be interpreted as a rank condition. It basically states that using a value i of the instrument, we can yield a binary instrument Z_i whose effect is truly distinct from those we can produce using the first $d + 1$ values of Z. A necessary condition for Assumption S3 to hold is that $K \geq d+2$, which is logical since full identification was obtained in the univariate case with $K \geq 3$. Theorem S1 shows that the model is identified under this condition. Its proof relies on a characterization of additive subgroups of $\mathbb{R}^d$, which can be found for instance in Bourbaki (1974).

Theorem S1 *If Equation (2.1) and Assumptions 1, S1- S3 hold, g is identified.*

2.2 The nonfree case

Without freeness, we can still use fixed points to achieve identification. However another element comes into play, namely the attractiveness of these fixed points. Attractiveness is not an issue in the univariate case since the functions are strictly increasing. Any fixed point of s can be reached by applying several times either s or s^{-1} and g is thus identified at the fixed point.

This is not true anymore in a multidimensional setting, as illustrated in Figure S1. Consider the bivariate case with $K = 2$, and let $x_f = (x_{1,f}, x_{2,f})$ denote a fixed point of $s_{12} = (s_{1,12}, s_{2,12})$. Suppose first that $s_{1,12}(x_1) > x_1$ if and only if $x_1 < x_{1,f}$, while $s_{2,12}(x_2) < x_2$ if and only if $x_2 < x_{2,f}$ (see Figure 1, case (a)). No sequence $(s_{12}^k(x))_{k\in\mathbb{N}}$ converges in $\mathcal{X}$. When $x = (x_1, x_2) \in (-\infty, x_{1,f}) \times (-\infty, x_{2,f})$, for instance, the sequence $(s_{1,12}^k(x_1))_{k\in\mathbb{N}}$ converges to

$x_{1,f}$ but the sequence $(s_{2,12}^k(x_2))_{k\in\mathbb{N}}$ tends to $-\infty$, with $(x_{1,f}, -\infty) \notin \mathcal{X}$. On the other hand, suppose that $s_{m,12}(x_m) < x_m$ if and only if $x_m < x_{m,f}$, for $m \in \{1, 2\}$ (Figure 1 case (b)). For any $x = (x_1, x_2)$, the sequence $(s_{12}^{-k}(x))_{k\in\mathbb{N}}$ converges to x_f.

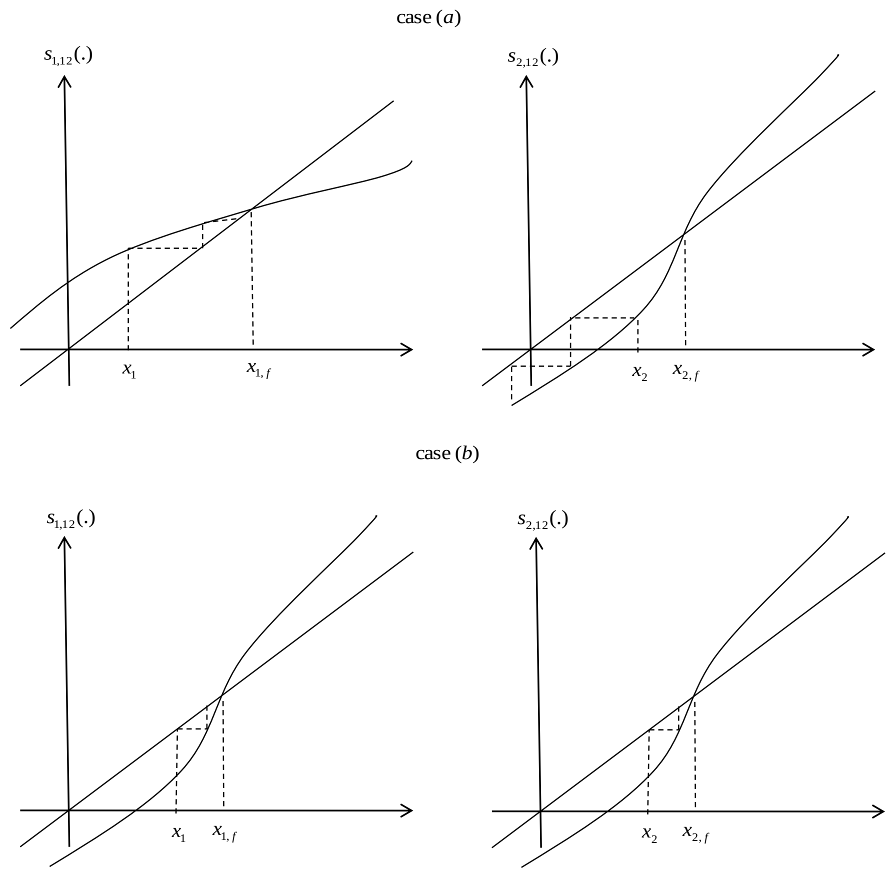

*Figure S*1: Illustration of the attractiveness issue under nonfreeness.

In short, a condition on the position of the coordinates of s_{12} is necessary and sufficient to secure identification when $K = d = 2$. The sufficiency part of this result actually extends to any K and d, as Proposition S2 shows.

Proposition S2 *Under Assumptions 1 and S1-S2, if there exists $s = (s_1, ..., s_d) \in \mathcal{S}$ with exactly one fixed point $x_f = (x_{1,f}, ..., x_{d,f})$ and such that for all $x =$*

$(x_1, ..., x_d)$, $sgn\left[(s_m(x) - x_m)(x_m - x_{m,f})\right]$ does not depend on $m \in \{1, ..., d\}$, then g is identified.

Even if the attractiveness condition may seem restrictive, it is important to note that only one function in the group has to satisfy this condition. Hence, it may hold even when no function s_{ij} admits an attractive fixed point, because we also have at hand all the compositions of the s_{ij}. To illustrate this idea, consider the generalized location-scale models of the form

$$X_m = \mu_m \left(\nu_m(Z) + \sigma_m(Z)\eta_m\right), \tag{2.2}$$

with $\sigma_m(Z) > 0$ and μ_m a strictly increasing and continuous function. Without loss of generality, we set $\sigma_1(1) = ... = \sigma_d(1) = 1$. Unless $\sigma_m(.)$ is constant for some m, all the functions s_{ij} admit a unique fixed point, which is not attractive in general. Nevertheless, under a simple rank condition, the model is identified because one can always construct a function $s \in \mathcal{S}$ with an attractive fixed point.

Theorem S2 *If Equation (2.2) and Assumptions 1, S1- S2 hold, and the rank of the matrix of typical (i, j) element $\ln \sigma_i(j + 1)$ is d, there exists $s \in \mathcal{S}$ that admits a unique and attractive fixed point. Thus, g is identified.*

3 Additional proofs

We begin by stating and proving Lemma S1, used in the proof of Theorem 2.

Lemma S1 *Suppose that s_{12} is C^2 with $s_{12}(x) - x > 0$ for all $x \in \mathcal{X}$. Then for any $a \in \mathcal{X}$, there exists an increasing C^2 diffeomorphism $\widetilde{r}$ from $[0, 1)$ to $[a, s_{12}(a))$ satisfying $\widetilde{r}(0) = a$, $\lim_{x\to 1} \widetilde{r}(x) = s_{12}(a)$, $\lim_{x\to 1} \widetilde{r}'(x) = \left[s_{12} \circ \widetilde{r}\right]'(0)$ and $\lim_{x\to 1} \widetilde{r}''(x) = \left[s_{12} \circ \widetilde{r}\right]''(0)$.*

Proof: we actually prove the stronger result that for any $(a, b, c) \in \mathcal{X} \times \mathbb{R}^{*+} \times \mathbb{R}$, there exists an increasing C^2 diffeomorphism $\widetilde{r}$ from $[0, 1)$ to $[a, s_{12}(a))$

satisfying $\widetilde{r}(0) = a$, $\lim_{x\to 1}\widetilde{r}(x) = s_{12}(a)$, $\widetilde{r}'(0) = b$, $\lim_{x\to 1}\widetilde{r}'(x) = s_{12}'(a)b$, $\widetilde{r}''(0) = c$ and $\lim_{x\to 1}\widetilde{r}''(x) = s_{12}''(a)b^2 + s_{12}'(a)c$. For that, we construct $\widetilde{r}'$ satisfying all the restrictions. We first define the functions $g_{d,e}$, for any $d > 0$ and $e \in (0, 1/4)$, as follows:

- on $[0, e)$, let $g_{d,e}(x) = b + cx(1 - x/2e)$;
- on $[e, 2e)$, $g_{d,e}(x) = 1/2[(b+ce/2-d)\sin(\pi(x-e)/e+\pi/2)+(b+ce/2+d)]$;
- on $[2e, 1-2e)$, $g_{d,e}(x) = d$;
- on $[1-2e, 1-e)$, $g_{d,e}(x) = 1/2[(f(e) - d)\sin(\pi(x - 1 + e)/e + \pi/2) + (f(e) + d)]$, with $f(e) = s_{12}'(a)b - (s_{12}''(a)b^2 + s_{12}'(a)c)e/2$;
- on $[1-e, 1)$, $g_{d,e}(x) = s_{12}'(a)b+[s"_{12}(a)b^2+s_{12}'(a)c](x-1)\,(1 + (x - 1)/2e)$.

By construction, $g_{d,e}$ and $g_{d,e}'$ are continuous. If e is small enough, $b + ce/2 > 0$ $f(e) > 0$ and $g_{d,e}(x) > 0$ for all $x \in [0, 1)$. Moreover, $g_{d,e}(0) = b$, $\lim_{x\to 1} g_{d,e}(x) = s_{12}'(a)b$, $g_{d,e}'(0) = c$ and $\lim_{x\to 1} g_{d,e}'(x) = s"_{12}(a)b^2 + s_{12}'(a)c$. Moreover, because $\lim_{\substack{d\to 0\\ e\to 0}} \int_0^1 g_{d,e}(x)dx = 0$ and for any $e \in (0, 1/4)$, $\lim_{d\to\infty} \int_0^1 g_{d,e}(x)dx = +\infty$, there exists, by the intermediate value theorem, (d^*, e^*) such that $f(e^*) > 0$ and $\int_0^1 g_{d^*,e^*}(x)dx = s_{12}(a) - a$. By construction, $\widetilde{r}(x) = a + \int_0^x g_{d^*,e^*}(t)dt$ satisfies all the restrictions of the lemma. ■

3.1 Proof of Proposition S1

Suppose first that $\sigma(Z)$ is constant, equal to one without loss of generality. We have $h^{-1}(i, x) = -\nu(i) + \mu^{-1}(x)$. As a result, $s_{ij}(x) = \mu(\nu(j) - \nu(i) + \mu^{-1}(x))$. For any $s \in \mathcal{S}$, there exists $(i_1, j_1, ..., i_p, j_p) \in \{1, ..., K\}^{2p}$ such that $s = s_{i_1 j_1} \circ ... \circ s_{i_p j_p}$. By a straightforward induction, $s(x) = \mu\left(\sum_{i=1}^K \nu(i)n(i) + \mu^{-1}(x)\right)$, where $n(i) = \sum_{l=1}^p \mathbb{1}\{j_l = i\} - \mathbb{1}\{i_l = i\}$. $s(x) = x$ for some x thus implies that $\sum_{i=1}^K \nu(i)n(i) = 0$, implying in turn that s is the identity function. Thus freeness holds. Now, suppose that $\sigma(Z)$ is not constant, and let i, j be such that $\sigma(i) \neq \sigma(j)$. We have

$$s_{ij}(x) = \mu\left[\nu(j) + \sigma(j)\frac{\mu^{-1}(x) - \nu(i)}{\sigma(i)}\right].$$

Thus, s_{ij} is different from the identity function and we can easily see that it admits a unique fixed point. Therefore, the nonfreeness property holds. $\blacksquare$

3.2 Proof of Theorem S1

As in the univariate case, we prove that any orbit is dense. The functions $s \in \mathcal{S}$ take the form

$$s(x_1, ..., x_d) = \left(\mu_1 \left[\sum_{k=2}^{K} n(k)\nu_1(k) + \mu_1^{-1}(x_1) \right], ..., \mu_d \left[\sum_{k=2}^{K} n(k)\nu_d(k) + \mu_d^{-1}(x_d) \right] \right),$$

for some $n = (n(2), ..., n(K)) \in \mathbb{Z}^{K-1}$. Moreover, any $n \in \mathbb{Z}^{K-1}$ corresponds to a function $s \in \mathcal{S}$. We thus have

$$\mathcal{O}_{x_0} = \left\{ \left(\mu_1 \left[\sum_{k=2}^{K} n(k)\nu_1(k) + \mu_1^{-1}(x_{01}) \right], ..., \mu_d \left[\sum_{k=2}^{K} n(k)\nu_d(k) + \mu_d^{-1}(x_{0d}) \right] \right), \right.$$
$$\left. (n(2), ..., n(K)) \in \mathbb{Z}^{K-1} \right\}.$$

By continuity of $\mu_1, ..., \mu_d$, it suffices to show that $H = \{\sum_{k=2}^{K} n(k)A'_{k-1},\ n(k) \in \mathbb{Z}\}$ is dense in $\mathbb{R}^d$. Because H is an additive subgroup of $\mathbb{R}^d$, it suffices to show (see, e.g., Bourbaki, 1974, paragraph 1, n°3) that

$$< H, x > \subset \mathbb{Z} \Longrightarrow\ x = 0, \tag{3.1}$$

where for any $x \in \mathbb{R}^d$,

$$< H, x >= \{h'x,\ h \in H\} = \left\{ \sum_{k=2}^{K} n(k)A_{k-1}x,\ n(k) \in \mathbb{Z} \right\}.$$

Suppose that $< x, H > \subset \mathbb{Z}$ for some $x \in \mathbb{R}^d$. Then $A_k x \in \mathbb{Z}$ for all $k = 1, ..., d$. Choosing $i > d + 1$ as in Assumption S3, we also have $A_i x \in \mathbb{Z}$. This implies that $\sum_{k=1}^{d} \lambda_k(A_k x) \in \mathbb{Z}$. Because $A_k x \in \mathbb{Z}$, $A_k x = 0$ for $k = 1, ..., d$ by Assumption S3. Because $(A_1, ..., A_d)$ are linearly independent, $x = 0$, implying (3.1). $\blacksquare$

3.3 Proof of Proposition S2

Suppose without loss of generality that $\operatorname{sgn}[(s_m(x_m) - x_m)(x_m - x_{m,f})] = -1$ for all $m = 1...d$. To prove Theorem S2, it suffices to show that $x_f = \lim_{k\to\infty} s^k(x)$ for all $x = (x_1, ..., x_d) \in \mathcal{X}$, or, equivalently, that for all $m = 1...d$, $x_{m,f} = \lim_{k\to\infty} s_m^k(x_m)$. If $x_m < x_{m,f}$, a straightforward induction shows that $(s_m^k(x))_{k\in\mathbb{N}}$ is increasing and bounded above by $x_{m,f}$. Because s has a unique fixed point, $x_{m,f} = \lim_{k\to\infty} s_m^k(x_m)$. Similarly, if $x_m > x_{m,f}$, $s_m^k(x)$ is decreasing and bounded below by $x_{m,f}$, and the sequence also converges to $x_{m,f}$. $\blacksquare$

3.4 Proof of Theorem S2

First, some algebra shows that functions $s \in \mathcal{S}$ take the form

$$s(x_1, ..., x_d) = \left(\mu_1\left[\alpha_1 + \left(\prod_{k=2}^{K}\sigma_1(k)^{e_k}\right)\mu_1^{-1}(x_1)\right], ..., \right.$$
$$\left. \mu_d\left[\alpha_d + \left(\prod_{k=2}^{K}\sigma_d(k)^{e_k}\right)\mu_d^{-1}(x_d)\right]\right),$$

for some $(\alpha_1, ..., \alpha_d) \in \mathbb{R}^d$ and $(e_2, ..., e_K) \in \mathbb{Z}^{K-1}$. Moreover, any $e \in \mathbb{Z}^{K-1}$ corresponds to a function $s \in \mathcal{S}$. Noting $\beta_m = \prod_{k=2}^{K}\sigma_m(k)^{e_k}$, the function s admits a unique attractive fixed point $x_f = (x_{1,f}, ..., x_{d,f})$ if, for all m, $0 < \beta_m < 1$. Indeed $\mu_m(\alpha_m + \beta_m\mu_m^{-1}(x_{m,f})) = x_{m,f}$ if and only if $\mu_m^{-1}(x_{m,f}) = \frac{\alpha_m}{1-\beta_m}$. Moreover, $\mu_m(\alpha + \beta\mu_m^{-1}(x_m)) > x_m$ for $x_m < x_{m,f}$. Thus, by Proposition S2, it suffices to show that there exists $(e_2, ..., e_K) \in \mathbb{Z}^{K-1}$ such that

$$\left(\prod_{k=2}^{K}\sigma_m(k)^{e_k}\right) < 1 \text{ for all } m \in \{1, ..., d\}. \tag{3.2}$$

Let M denote the $d \times K - 1$ matrix of typical (i, j) element $\ln\sigma_i(j + 1)$. Because M is full rank by assumption, there exists $u \in \mathbb{R}^{K-1}$ such that $Mu = -(1, ..., 1)'$. Thus, by density of $\mathbb{Q}^{K-1}$, there exists $\widetilde{u} \in \mathbb{Q}^{K-1}$ such that $M\widetilde{u} < 0$, where the inequality should be understood componentwise. Moreover, $\widetilde{u}$

can be written $(e_2/D, ..., e_K/D)'$, where $(e_2, ...e_K, D) \in \mathbb{Z}^K$. This implies that $M(e_2, ...e_K)' < 0$ which is equivalent to (3.2). $\blacksquare$

References

Bourbaki, N. (1974), *Topologie générale, chap. VII*, Hermann, Paris.

Guerre, E., Perrigne, I. & Vuong, Q. (2009), 'Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions', *Econometrica* **77**, 1193–1227.

Torgovitsky, A. (2014), 'Identification of nonseparable models using instruments with small support', *Econometrica* **Forthcoming**.