Identification of Mixture Models Using Support Variations*

Xavier D'Haultfoeuille † Philippe Février‡

March 2010

Abstract

We consider the issue of identifying nonparametrically mixture models. In these models, all observed variables depend on a common and unobserved component, but are mutually independent conditional on it. Such models are important in the measurement error, auction and matching literatures. Traditional approaches rely on parametric assumptions or strong functional restrictions. We show that these models are actually identified nonparametrically if a moving support assumption is satisfied. More precisely, we suppose that the supports of the observed variables move with the true value of the unobserved component. We show that this assumption is theoretically grounded, empirically relevant and testable. Finally, we compare our approach with the diagonalization technique introduced by Hu and Schennach (2008), which allows to obtain similar results.

Keywords: mixture models, nonparametric identification, measurement error, auctions, matching.

JEL classification numbers: C14, D44.

*We would like to thank Jean-Marc Robin, Bernard Salanié and Susanne Schennach for their useful comments.

†CREST (INSEE). E-mail address: xavier.dhaultfoeuille@ensae.fr.

‡CREST (ENSAE). E-mail address: fevrier@ensae.fr.

1 Introduction

In this paper, we consider nonparametric mixture models where all observed variables depend on a common and unobserved component, but are mutually independent conditional on it. Such models have important applications in economics. The main one is probably the measurement error model, in which extensive attention has been devoted to identifying the effect of an unobserved variable when only measures of it are available. While the literature on this topic is vast (see, e.g., Carroll et al., 2006 for a survey), most of the papers focus on the case of classical measurement errors, for which errors are either independent of the mismeasured variable or have a zero mean conditional on it (see, e.g., Hausman et al., 1991, Li, 2002, Schennach, 2004 and 2007). Yet, this assumption is likely to fail in many context (see, e.g., Bound and Krueger, 1991). Building on the ideas of Hu (2008), Hu and Schennach (2008) explain in a recent paper how to recover the effect of the true variable in the general case of nonclassical measurement errors with continuous variable. Under an injectivity condition one integral operators, they show that identification can be achieved through an eigenvalue-eigenfunction decomposition. This method has also been useful to answer other economic questions in which mixture models are present. Hu and Shum (2009*a*) and Hu et al. (2009) have relied on it to study respectively entry and heterogeneity in auctions, whereas Hu and Shum (2009*b*) have applied the same technique to dynamic models with unobserved state variables. Finally, another example where mixtures models apply is the matching literature. In this case, we observe equilibrium outcomes from the matches between heterogeneous agents. The aim is to recover the link between the unobserved heterogeneity of the individuals and the outcome of the matches.

We propose here an alternative approach to Hu and Schennach (2008) and introduce a very simple sufficient condition for the model to be identified. More precisely, we suppose that the observed variables have a compact support that moves with the unobserved variable. When this "moving support assumption" is satisfied, and a necessary normalization is imposed, the model is identified without any other restriction. This approach complements Hu an Schennach's one in the sense that for some models, our condition is satisfied while theirs fails to hold, and conversely.

We believe that our identification result is interesting for several reasons. First, the "moving assumption" is naturally satisfied in different economic models. This is for example the case in the matching literature. Building on Becker (1973) result, Shimer and Smith (2000) derive sufficient conditions to extend assortative matching in an environment with search frictions. In this model, at equilibrium, workers match with firms of different qualities. As seen below in Figure 1 (also Figure 1 in Shimer and Smith's paper), the set of firms with which a worker can match is increasing in the own quality of the worker and the "moving support" assumption is satisfied. Similarly, in an auction model with a reserve price unobserved by the econometrician, both the lower and upper bounds of the bids vary with the unobserved reserve price (Riley and Samuelson, 1981).

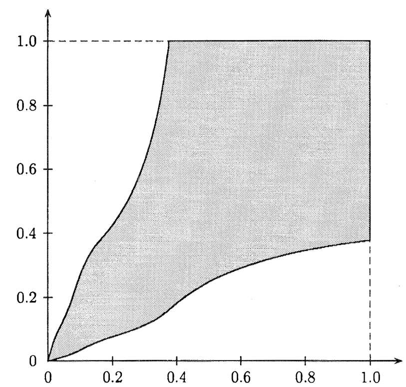

Figure 1: An example of matching set with search friction (taken from Shimer and Smith, 2000).

Second, our assumption is easy to interpret economically. In the measurement error model, the underlying idea is that the mismeasured variable cannot be too far from the true value of the variable. Actually, in some cases, both the measurement and the true variable are observed. Such data, even if unusual, are interesting to check directly if the "moving support assumption" is reasonable or not. The validation sample of the PSID analyzed by Rodgers et al. (1993), for instance, seems to support our condition (see Figure2).[1] Finally, even when the true value of the variable is unknown, we show that it is possible to test the "moving support assumption", using results from the statistical literature on extreme values (see, e.g., Embrechts et al., 1997).

[1] In this figure, as well as in the rest of the paper, the frontier functions have been estimated by the DEA estimator (see, e.g., Farrell, 1957).

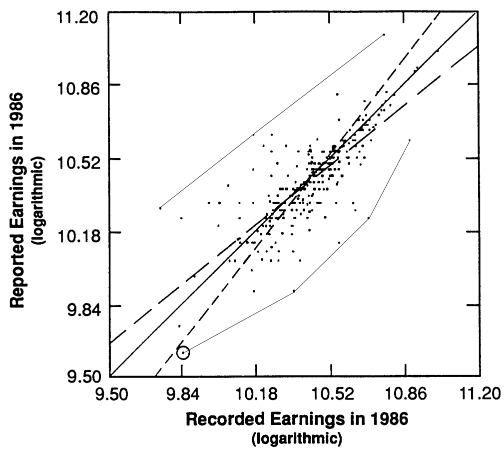

Figure 2: Minimum and maximum values of reported earnings conditional on true earnings (taken from Rodgers et al., 1993).

The paper is organized as follows. Section 2 presents the model and our main identification result. The theoretical ground, empirical relevance and testability of our "moving support assumption" is discussed in Section 3. Section 4 is devoted to some extensions and to the comparison with Hu and Schennach's framework. Section 5 concludes. All proofs are deferred to the appendix.

2 The model and main result

We define in this section the general mixture model we focus on. We consider K real random variables $(X_1, ...X_K)$ which are observed by the econometrician. All depend on a real continuous variable X^*, which is unobserved. We suppose, without loss of generality, that X^* is uniformly distributed.[2] The aim of the econometrician is to recover the distribution of X_k conditional on X^*.

The first assumption defines the mixture structure.

Assumption 1 *$K \geq 3$ and $(X_1, ...X_K)$ are independent conditional on X^*.*

[2]Other normalizations are possible. We could impose, similarly to Hu and Schennach (see their Assumption 5), that the mode, the median or some quantile of the measurement error is equal to zero. We come back to this question in Subsection 4.2.

This framework is well suited to various economic settings, such as models with measurement errors on a covariate. Typically, we seek to measure the effect of a variable X^* on an outcome $Y(= X_3) = f(X^*, \nu)$ but only observe two variables (X_1, X_2) related to X^*. These variables may represent two measures of X^*, so that $X_k = \varphi_k(X^*, \eta_k)$, $k \in \{1, 2\}$ (see, e.g., Hausman et al., 1991 or Schennach, 2004, for papers studying models with repeated measures). Alternatively, we may observe only one measure X_1 and an instrument $Z(= X_2)$ of X^* such that $X^* = \psi(Z, \xi)$ (see, e.g., Newey, 2001 or Schennach, 2007 for studies of instrumental models with measurement errors). In the first case, Assumption 1 is satisfied if (ν, η_1, η_2) are independent conditional on X^*, while it holds in the second if (Z, ν, η_1, ξ) are independent. Assumption 1 is equivalent to Assumption 2 of Hu and Schennach (2008), so that our framework is identical to theirs.

As mentioned in the introduction, this setting is however more general than the measurement error model and applies to several other economic frameworks. Auctions with unobserved heterogeneity is one example. Let us indeed consider a good which is sold by an auction mechanism. This good has a characteristic X^* which is observed by the K bidders and affects their valuation $(V_1, ..., V_K)$. Conditional on X^*, $(V_1, ..., V_K)$ are independent, but may be non identically distributed if bidders are asymmetric. The econometrician observes the bids $B_k(= X_k) = b_k(V_k, X^*)$ but neither $(V_1, ..., V_K)$ nor X^*. In such a case, $(B_1, ..., B_K)$ are independent conditional on X^*. The ultimate goal in this literature is to recover the distribution of V_k conditional on X^*. However, in general, the function b_k is known by the theory and it is thus sufficient to recover the distribution of B_k conditionally on X^*. Such auction models with unobserved heterogeneity have been studied recently by Krasnokutskaya (2009) and Hu et al. (2009), the latter applying Hu's (2008) methodology. Common value models also fit this framework, the unobserved variable being the true unknown value of the good. D'Haultfoeuille and Février (2009) study this model using the methodology presented in this paper. More generally, mixture models are useful as soon as there are unobserved components and/or unobserved heterogeneity. This is generally the case in panel data or in dynamic models with unobserved state variable (Hu and Shum, 2009*b*).

We now introduce two other assumptions. First, letting $F_{X_k|X^*}(.|.)$ denote the cumulative distribution function of X_k conditional on X^*, we assume the following mild regularity conditions.

Assumption 2 *For all $k \in \{1, ..., K\}$ and $u \in [0, 1]$, the support of X_k conditional on $X^* = u$ is an interval and we denote it by $[\underline{X}_k(u), \overline{X}_k(u)]$. Moreover, $x \mapsto F_{X_k|X^*}(x|u)$*

is continuously differentiable for all u, $u \mapsto F_{X_k|X^*}(x|u)$ *is continuous for all* x *and* $\underline{X}_k(.)$ *and* $\overline{X}_k(.)$ *are continuously differentiable.*

Second, we impose the "moving support assumption" that states that the supports moves with the true value of the unobserved variable. This is our main condition and we distinguish two cases depending on the number of variables that satisfy this assumption.

Assumption 3 *For* $k \in \{1, 2, 3\}$, $\underline{X}_k(.)$, $\underline{X}'_k(.) > 0$ *and* $\overline{X}'_k(.) > 0$.

Assumption 4 *For* $k \in \{1, 2\}$, $\underline{X}'_k(.) > 0$ *and* $\overline{X}'_k(.) > 0$.

The fact that the bounds of the support are increasing functions of X^* reflects the positive link between X^* and the Xs. Such a pattern is not very restrictive and is, for example, a consequence of the maximum likelihood ratio property that states that $\frac{f_{X_k|X^*}(x|x_1^*)}{f_{X_k|X^*}(x|x_0^*)}$ is an increasing function of x when $x_1^* > x_0^*$. Here, we need to reinforce this condition by stating that higher values of X^* lead to strictly higher values of X_k, i.e., that the observed variable has a support that moves with the true value of the unobserved component.

Under our assumptions, the model is identified.

Theorem 2.1 *Under Assumptions 1-3,* $f_{X_k|X^*}(.|.)$ *is identified for all* $k \in \{1, ..., K\}$. *Under Assumptions 1, 2 and 4,* $f_{X_k|X^*}(.|.)$ *is identified for all* $k \geq 3$.

The intuition behind the proof is the following. Suppose for simplicity that the model is symmetric and that for all k, $F_{X_k|X^*}(x|u) = F_{X|X^*}(x|u)$ on a support $[\underline{X}(X^*), \overline{X}(X^*)]$. As shown by Figure 3, the range of X^* compatible with an observation $X_1 = x$ is limited, and so is the range of X_2 that one can observe in the data when $X_1 = x$. More formally, observing $X_1 = x$, we know that X^* belongs to the set $[\overline{X}^{-1}(x), \underline{X}^{-1}(x)]$. Hence, X_2 belongs to the set $[\underline{X} \circ \overline{X}^{-1}(x), \overline{X} \circ \underline{X}^{-1}(x)]$ that we denote by $[\underline{S}(x), \overline{S}(x)]$.

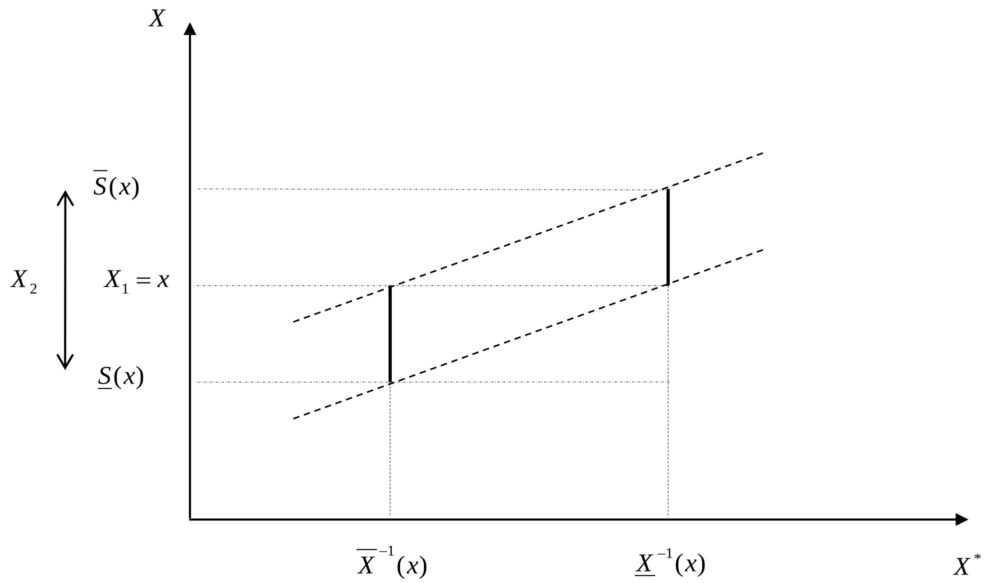

Figure 3: Minimum and maximum values of X_2 when $X_1 = x$

Hence, and as described in Figure 4, observing both the values $X_1 = x$ and $X_2 = \overline{S}(x)$, allows us to pin down the unique $X^* = \underline{X}^{-1}(x)$ compatible with these two values. When $K \geq 3$, it is then possible to identify, for all x, $f_{X|X^*}(.|\underline{X}^{-1}(x))$ by looking at the distribution of a third observation X_3 conditional on observing $X_1 = x$ and $X_2 = \overline{S}(x)$. Indeed, by the conditional independence assumption,[3]

$$f_{X_3|X_1,X_2}(.|x, \overline{S}(x)) = f_{X_3|X_1,X_2,X^*}(.|x, \overline{S}(x), \underline{X}^{-1}(x)) = f_{X|X^*}(.|\underline{X}^{-1}(x)).$$

[3]This equality is not rigorous because the density f_{X_1,X_2} is equal to zero at $(x, \overline{S}(x))$. To overcome this issue, we have to consider instead the event $(X_1, X_2) \in A_\delta(x_1) = [x_1 - \delta, x_1 + \delta] \times [\overline{S}(x_1 - \delta), \overline{S}(x_1 + \delta)]$, and let $\delta \to 0$. The formal proof is given in appendix.

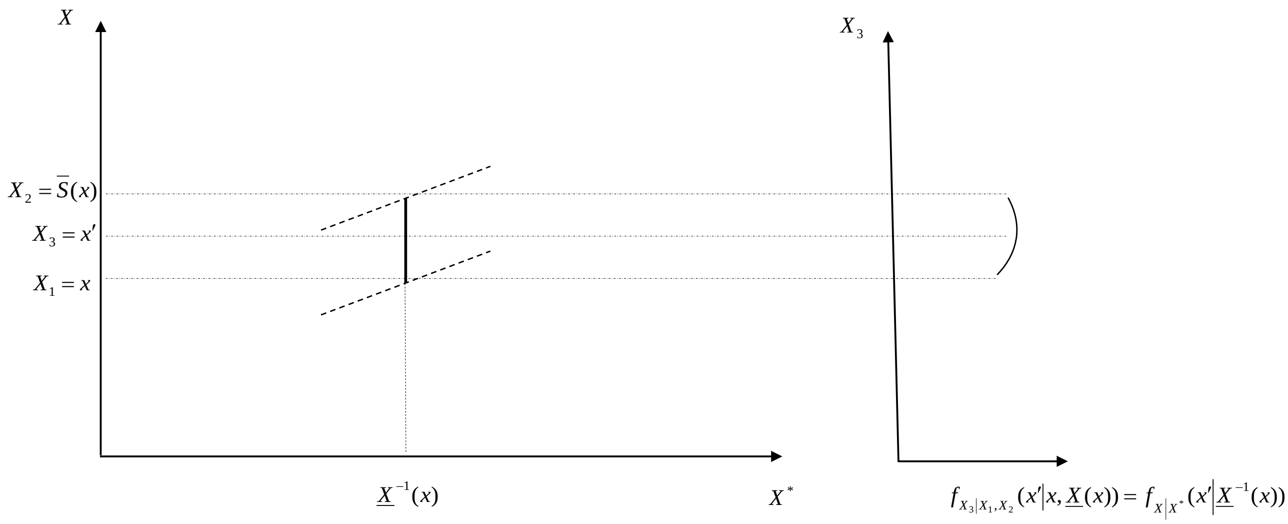

Figure 4: Intuition on the identification of $f_{X|X^*}$

In a last step, because X^* is uniformly distributed by our normalization, we recover the function $\underline{X}^{-1}(.)$ by using[4]

$$f_{X^*}(\underline{X}^{-1}(x)) = \underline{X}^{-1}(x) = \frac{f_{X_1,X_2}(x, \overline{S}(x))}{f_{X|X^*}(x|\underline{X}^{-1}(x))f_{X|X^*}(\overline{S}(x)|\underline{X}^{-1}(x))},$$

in which all functions in the right-hand side are identified. $\underline{X}(.)$ is thus identified and $f_{X|X^*}(.|.) = f_{X|X^*}(.|\underline{X}^{-1}(\underline{X}(.)))$ also is.

This informal proof also gives the intuition why having only two observed variables that satisfy the moving support assumption is sufficient to identify the conditional distribution of other observed variables. Two variables are sufficient to pin down the value of X^*, and the distribution of any other variable can thus be recovered. This result is particularly important for the measurement error framework. Indeed, no restriction has to be made on the dependent variable Y. Only the supports of two measures of the unobserved variable have to vary with the underlying unobserved true variable.

[4]Once again, this equation is not rigorous as the ratio may not be properly defined. The formal proof considers a limit reasoning to circumvent this issue.

3 The moving support assumption

3.1 Theoretical grounding

The crucial assumption for our result to be valid is the moving support assumption. It is thus interesting to look at some economic models in which our assumption is satisfied.

First, the moving assumption is fulfilled in the classical measurement error model or with multiplicative errors, as soon as the error term has a bounded support. Indeed, if $X_k = X^* + \eta_k$ (resp. $X_k = X^* \times \eta_k$) and $\eta_k \in [\underline{\eta}, \overline{\eta}]$, the support of X_k conditional on X^* is $[X^* + \underline{\eta}, X^* + \overline{\eta}]$ (resp. $[X^* \times \underline{\eta}, X^* \times \overline{\eta}]$) and changes with X^*. Such a model is also used by Krasnokutskaya (2009) to take into account unobserved heterogeneity in first price auctions.

The moving support assumption, or part of it, may also derive naturally from the data at our disposal. This is typically the case with truncated or censored data if the truncation or the censoring is unobserved. For example, the econometrician may observe several wages for an individual only if these wages are above his unobserved reserve wage. In such a case, almost by definition, the lower bound of the wages moves with the unobserved component i.e. the unobserved reserve wage : $\underline{X}(X^*) = X^*$. Let us also mention the tobacco example cited by Hu and Schennach (2008). If tobacco consumption is likely to be either truthfully reported or under-reported, the upper bound of the reported value corresponds to the true underlying consumption. At least one part of Assumption 3 is therefore satisfied because $\overline{X}(X^*) = X^*$.

More structurally, the moving support assumption may derive from the theoretical economic model underlying the production of the data. Let us consider for example an auction with reserve prices X^* which are unobserved by the econometrician. We suppose that N potential risk neutral and symmetric bidders with valuations $(V_1, ..., V_N) \in [\underline{V}, \overline{V}]^N$ participate to this auction. We note $F_V(.)$ (resp. $f_V(.)$) the distribution (resp. density) function of V_i. Finally, we suppose that before bidding, the bidders learn the number n of effective bidders i.e. the number of bidders with valuations greater than X^*. In such a case, the equilibrium function $b(.,.,.)$ is given by :

$$b(V, X^*, n) = V - \frac{\int_{X^*}^{V} F_V^{n-1}(u)du}{F_V^{n-1}(V)}$$

Hence, for all n, conditional on X^*, the observed bids[5] $(X_1, ..., X_n) = (b(V_1, X^*, n), ..., b(V_n, X^*, n))$ belong to the set $[X^*, b(\overline{V}, X^*, n)]^N$. Both bounds are strictly increasing in X^* and the moving support assumption is a consequence of the theoretical model.[6]

Similarly, let us consider the matching model with search frictions developed by Shimer and Smith (2000)[7]. In their model, they assume a continuum of heterogeneous agents. Two agents of type X^* and Y in $[0, 1]$ can match to produce $f(X^*, Y)$, where f is strictly increasing in both arguments. At each instant, agents are either matched or unmatched, and nature destroys any match with a positive probability. Unmatched agents constitute the pool of searchers that are trying to form new matches. Shimer and Smith characterize the equilibrium matching sets. They prove in particular that under regularity and supermodularity assumptions, positively assortative matching is ensured.[8] This, in turn, implies that the lower and upper bound functions of the matching set are nondecreasing,[9] as depicted in Figure 1. Let us then suppose that the econometrician observes several matches between, for example, firms and workers on the job market, and their corresponding matching outputs.[10] Hence, for a worker of unobserved type X^*, the econometrician observes $(X_1, ..., X_n) = (f(X^*, Y_1), ..., f(X^*, Y_n))$. Given Shimer and Smith (2000) results, the support of X_i is given by $[f(X^*, \underline{Y}(X^*)), f(X^*, \overline{Y}(X^*))]$. Because $x \mapsto f(x, y)$ is strictly increasing whereas $\underline{Y}(.)$ and $\overline{Y}(.)$ are nondecreasing, the moving support assumption is satisfied. Hence, the distribution of X_i conditional on X^* is identified.

3.2 Empirical evidence

More generally, and even if the moving support condition does not come directly from the model, a nice feature of our condition is that its definition relies only on the primitives of the model and is easy to interpret. Basically, it means that the mismeasured variable cannot be too far from the true variable. This assumption seems reasonable to us and we believe it should be satisfied in several empirical settings.

First, one can verify whether the moving assumption is satisfied or not when looking at

[5] We suppose here without loss of generality that bidders $1, ..., n$ have a valuation greater than X^*.

[6] This result holds regardless of whether $\overline{V} < +\infty$ or not, provided that $E(V) < +\infty$.

[7] We thank Jean-Marc Robin for suggesting us this example.

[8] For more details, see their Proposition 6.

[9] For more details, see their Proposition 3.

[10] Wages are usually observed, rather than the production itself. However, in simple models, there is a one to one relationship at equilibrium between the wages and the output.

some data for which an auxiliary dataset containing the true variables is available. This is typically the case for the validation sample taken from the PSID and used by Rodgers et al. (1993) to better understand how to model measurement errors. These data contain indeed both reported earnings and the exact earnings obtained from administrative records. It is thus interesting to see if Assumption 3 is relevant in these data. As mentioned in the introduction (see Figure 2), the minimum and maximum reported earnings appear to be strictly increasing functions of the true earnings, supporting our condition.

Such data are unfortunately quite unusual and one may want to have other ways to check the moving support assumption. In the case of unobserved heterogeneity, an indirect and informal test consists of looking at the link between the measured variable and some observed components. Indeed, if the observed components satisfy the moving support assumption, it is reasonable to think that the unobserved one also does. There is no reason, a priori, to treat differently the observed and unobserved components. Figure 5 for example shows the link between the highest bid P and an estimation P^* of the value of the wines given by the auctioneer in wine auctions at Drouot before the auctions take place (See Février et al., 2005, for more details). It clearly appears that the support of P strictly moves with P^*. Other observed components that influence positively the bids (quality of the wine, quality of the year,...) display similar patterns. The bids strictly increase when some positive information is revealed about the wines. Hence, if such a pattern is at stake for the observed components, it indicates indirectly that this should also be the case for the unobserved ones.

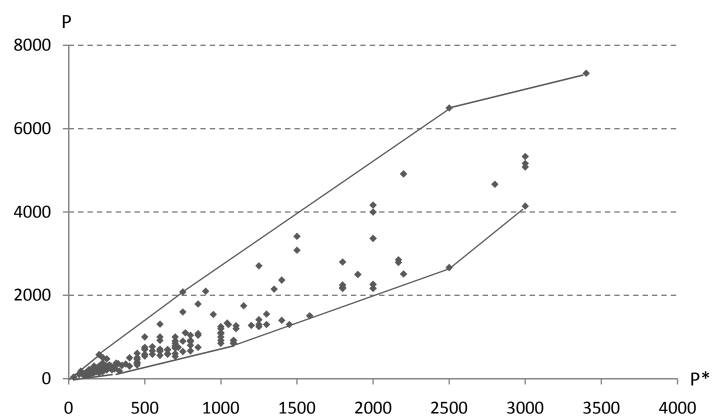

Figure 5: Values of the winning bids P given the estimation P^* of the auctioneer

Finally, it is possible to verify directly if the moving assumption is reasonable or not. Indeed, as already mentioned, we know that the set of values X_2 that one can observe with $X_1 = x$ is given by $[\underline{S}(x), \overline{S}(x)]$. Figure 6 thus shows that we are able to recover the upper and lower bounds using only the data. As these functions are strictly increasing on $[\overline{X}(0), \underline{X}(1)]$ if and only if the moving support assumption is satisfied, it is possible to check directly our main assumption by constructing empirically these functions.[11] As an illustration, Figure 7 displays the dependence between several test scores of French students from the 1997 panel of the French Ministry of Education.[12] Such test scores are often used as proxies for the unobserved cognitive ability X^*, but typically suffer from the measurement error problem. The left graph corresponds to the dependence between first and third grade test scores (X_1 and X_2), while the right one presents the sixth grade maths score as a function of the sixth grade French score (X_1' and X_2'). In both cases, the supports of these test scores seem to be strictly increasing, providing evidence that the moving support assumption is satisfied.

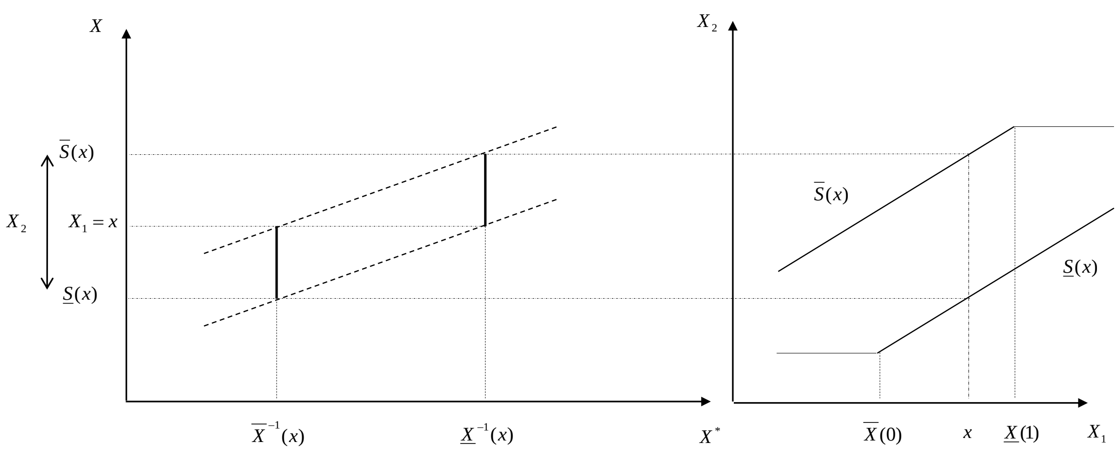

Figure 6: Minimum and maximum values of X_2 when $X_1 = x$

[11]Note that it is possible to identify $\overline{X}(0)$ and $\underline{X}(1)$, as shown in the proof of Theorem 2.1.
[12]We restrict ourselves to students who did not repeat any grade before the sixth grade.

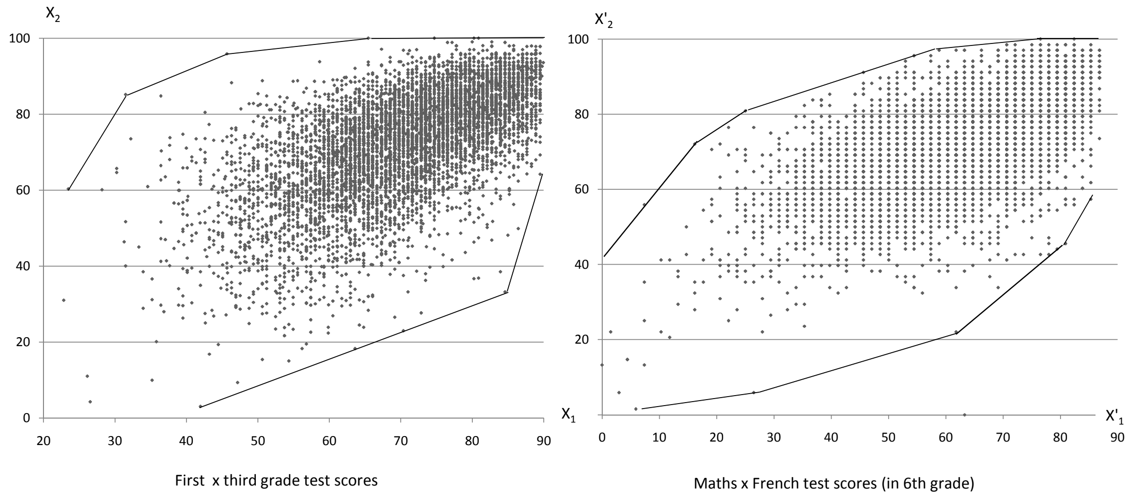

Figure 7: The moving support assumption for test scores of French students

3.3 Formal tests

The previous examples make us believe that the moving support assumption is theoretically and empirically relevant. Nevertheless, a potential concern is that the observed patterns could stem from the strict monotonicity of all quantiles, except the minimum and maximum. This would typically be the case if $X_k = X^* + \nu_k$, with ν_k following a normal or student distribution. It is thus desirable to formally test for the moving support assumption.

Assumption 4 imposes two restrictions on the distributions of X_1 and X_2 conditional on X^*.[13] First, their support should be compact. Second, the corresponding upper and lower bounds should be strictly increasing as functions of X^*. Such distributions are unobserved, but Figure 6 suggests that we can use instead the distribution of X_1 (resp. X_2) conditional on X_2 (resp. X_1). To do so, let $(A_j)_{j=1...J}$ and $(B_{j'})_{j'=1...J'}$ be two partitions of the supports of X_1 and X_2 into finite intervals. Then one can show that X_1 and X_2 have finite supports conditional on X^* if and only if, for all (j, j'), the supports of X_2 conditional on $X_1 \in A_j$ and of X_1 conditional on $X_2 \in B_j$ are finite. Similarly, if A_2 contains larger values than A_1, one can test for the strict monotonicity condition by testing that the upper bound $\overline{X}_{2j}$ ($j \in \{1, 2\}$) of the support of X_2 conditional on $X_1 \in A_j$ is strictly increasing in j. We

[13]The tests would be similar for Assumption 3.

investigate below how both points can be formally tested, under the standard condition that we observe a sample $(X_{1i}, ..., X_{Ki})_{i=1...n}$ of independent and identically distributed variables.

To test for the compact support assumption, let us consider a given partition $(A_j)_{j=1...J}$ of the support of X_1. Our purpose is to test, for all j, that the supremum $\overline{X}_{2j}$ of the support of X_2 conditional on $X_1 \in A_j$ is finite (the test for a finite infimum is similar). We actually test for the slightly stronger condition that $F_{X_2|X_1 \in A_j} \in \mathcal{P}$, where $\mathcal{P}$ is the subset of the set $\mathcal{F}$ of all distributions functions defined by[14]

$$\begin{aligned} \mathcal{P} &= \{F \in \mathcal{F} / \sup(\text{support}(F)) = \overline{x}_F < \infty \text{ and } F(x) = 1 - (\overline{x}_F - x)^{\alpha} L(1/(\overline{x}_F - x)) \\ &\quad \text{for some } \alpha > 0 \text{ and some slowly varying function } L\}. \end{aligned}$$

We rely on the following result (see, e.g., Theorem 3.3.12 of Embrechts et al., 1997).

Theorem 3.1 *$F_{X_2|X_1 \in A_j} \in \mathcal{P}$ if and only if $F_{X_2|X_1 \in A_j}$ belongs to the domain of attraction of the Weibull distribution.*

This result is related to the extreme value theory that shows that under mild restrictions, the maximum of an i.i.d. sample drawn from F converges in distribution either to a Weibull, Gumbel or Fréchet distribution. Theorem 3.1 shows that the Weibull domain of attraction only encompasses distributions F with finite supremum. This literature also shows that distributions F with light tails such as the normal or exponential ones converge to a Gumbel, whereas distributions F with heavy tails such as the Pareto or Cauchy ones converge to a Fréchet. Actually, the three limit distributions can be gathered together in the family of generalized extreme value distributions, indexed by a parameter $\xi \in \mathbb{R}$. ξ is negative for the Weibull, zero for the Gumbel and positive for the Fréchet. Moreover, for each distribution function F, there is a unique ξ_F that corresponds to the limit in distribution of the maximum of the sample. Hence, testing for $F_{X_2|X_1 \in A_j} \in \mathcal{P}$, that is basically that the supremum of the support is finite, is equivalent to testing for $\xi_{F_{X_2|X_1 \in A_j}} < 0$. This equivalence is fortunate because inference on this parameter can be led in many ways, using for instance the Pickland estimator or modifications of it (see Embrechts et al., 1997, Chapter 6). Hence, implementing such a test for several A_j allows us to formally test whether the data are compatible with the compact support assumption.

We now turn to the test of the strict monotonicity of the upper bound, under the maintained assumption that supports are compact. More precisely, we test if $\overline{X}_{21} < \overline{X}_{22}$ with $\overline{X}_{2j}$ the

[14]Recall that a real function L is slowly varying if for all $t > 0$, $\lim_{x \to \infty} L(tx)/L(x) = 1$.

upper bound of the support of X_2 conditional on $X_1 \in A_j$. We suppose that $F_{X_2|X_1\in A_j}$ belongs to the subset $\mathcal{P}'$ of $\mathcal{P}$ defined by

$$\mathcal{P}' = \left\{F \in \mathcal{P}/\exists m > 0/F^{(1)}(\overline{x}_F) = ... = F^{(m-1)}(\overline{x}_F) = 0 \text{ and } F^{(m)}(\overline{x}_F) \neq 0\right\}.$$

Using a Taylor expansion of $F_{X_2|X_1\in A_j}$ at $\overline{X}_{2j}$, one can show that

$$n^{1/m}(\overline{X}_{2j} - M_{nj}) \xrightarrow{L} \mathcal{W}\left(\frac{(-1)^{m-1}P(X_1 \in A_j)F^{(m)}_{X_2|X_1\in A_j}(\overline{X}_{2j})}{m!}, m\right),$$

where M_{nj} is the maximum of the $(X_{2i})_{i:X_{1i}\in A_j}$ and $\mathcal{W}$ denotes the family of Weibull distributions. Hence we can consistently test for $\overline{X}_{21} = \overline{X}_{22}$ against $\overline{X}_{21} < \overline{X}_{22}$ since, under the null, the distribution of $T = n^{1/m}(M_{n2} - M_{n1})$ is identified, whereas T tends to infinity under the alternative. Such a test could be implemented with different intervals A_j. The rejection of all these tests strongly supports the strict monotonicity condition.

4 Discussion and extensions

4.1 Fewer variation in the support

Up to now, we have supposed that both the lower and upper bound are strictly increasing functions. One may wonder whether our result would still hold in general with fewer variation in the support. We show that it is possible to extend Theorem 2.1, in the symmetric case where the X_k are identically distributed conditional on X^*, to situations where only one bound is strictly increasing.

Assumption 5 *The $(X_k)_{k=1...K}$ are identically distributed conditional on X^*, $\overline{X}'(.) > 0$, $\underline{X}(.) = \underline{X}$ is constant and $F_{X|X^*} \in \mathcal{P}'$.*

Theorem 4.1 *Under Assumptions 1, 2 and 5, $f_{X|X^*}(.|.)$ is identified .*

We also prove that without any variation in the support, the model is not identified in general. To the best of our knowledge, this is the first non identification result on these models. It indicates that restrictions, such as our moving support assumption or the injectivity assumption of Hu and Schennach (2008), are necessary to identify nonparametric mixture models. Quite surprisingly, this result does not depend on the number K of variables that we observe.

Theorem 4.2 *Under Assumptions 1 and 2 only, the distributions $f_{X_k|X^*}$ are not identified in general.*

4.2 Comparison with the diagonalization approach

Whereas Hu and Schennach (2008) rely on the injectivity of integral operators (see their Assumption 3), we mostly use the moving support condition. We believe that the merits of our condition are its simple economic meaning and its testable implications. On the contrary, no empirical test has been proposed on the injectivity condition yet. Even theoretically, not much is known about this condition. It is closely related to the completeness condition used in additive instrumental nonparametric models to secure identification. This latter condition holds in exponential models (see Newey and Powell, 2003), or in nonlinear models under an additive decomposition and a large support condition, but under restrictive technical conditions (see D'Haultfœuille, 2009). No result has been obtained otherwise. It is thus difficult to define which restrictions on the dependence between X_k and X^* are implied by the injectivity condition.

Nevertheless, it is worth mentioning that the injectivity condition and the moving support assumption are more complementary than anything else, as the two examples below emphasize.

Example 1 *Consider the case where $X_k = X^* + U_k$, where $(X^*, U_1, ..., U_K)$ are independent and $U_k \sim \mathcal{N}(0, \sigma_k^2)$. Then the moving support assumption obviously fail. On the other hand, Assumption 3 of Hu and Schennach (2008) is satisfied. In other terms, the operators $T_k(g) = \int f_{X_k|X^*}(x|u)g(u)du$ are injective on the space of integrable functions. Indeed*

$$T(g)(x) = e^{-\frac{x^2}{2\sigma_k^2}} \int e^{-\frac{xu}{\sigma_k^2}} \left(e^{-\frac{u^2}{2\sigma_k^2}} g(u) \right) du.$$

T is the product of a positive function and the Laplace transform of $\exp(-\frac{u^2}{2\sigma_k^2})g(u)$. By injectivity of this transform (see, e.g., Bellamn and Roth, 1984), $T_k(g) = 0$ if and only if $g = 0$.

Example 2 *Suppose now that $X_k = X^* + U_k$, where $(X^*, U_1, ..., U_K)$ are independent, X^* has a compact support $[\underline{x}, \overline{x}]$ and $U_k \sim \mathcal{U}[0, \alpha_k]$. The moving support assumption is satisfied. On the other hand, Assumption 3 of Hu and Schennach (2008) fails to hold. Indeed, $T_k(g)(x) = \int_{x-\alpha_k}^{x} g(u)du$. As a result, any function which is periodic with period*

α_k *on* $[\underline{x} - \alpha_k, \overline{x}]$*, which equals zero elsewhere and such that* $\int_0^{\alpha_k} g(u)du = 0$ *satisfies* $T_k(g) = 0$.

Another difference between the two papers is the choice of the normalization on X^*. We choose here for simplicity to normalize X^* to have a uniform distribution, whereas the normalization is achieved by Hu and Schennach (2008) through a link between one of the X_k and X^* (see their Assumption 3). More precisely, they suppose that there exists a known functional M such that

$$M\left[f_{X_k|X^*}(.|x^*)\right] = x^*. \tag{4.1}$$

Such a functional may be for instance the mean, the mode, the median or any quantile of the distribution. The choice of the normalization is however innocuous for the identification results. Suppose for example, in our framework, that X^* satisfies Condition (4.1) rather than being uniformly distributed. Then, by Theorem 2.1, we can identify $f_{X_k|\widetilde{X}^*}$ where $\widetilde{X}^*$ is uniformly distributed and defined by $\widetilde{X}^* = R(X^*)$, R being a strictly increasing function. By (4.1),

$$M\left[f_{X_k|\widetilde{X}^*}(.|\widetilde{x}^*)\right] = M\left[f_{X_k|X^*}(.|R^{-1}\left(\widetilde{x}^*\right))\right] = R^{-1}\left(\widetilde{x}^*\right).$$

As a result, R^{-1} is identified, and so are $f_{X_k|X^*}$ and f_{X^*}.

5 Conclusion

This paper proposes an alternative and complementary approach to Hu and Schennach (2008)'s one to identify mixture models. Our result relies on a moving support assumption that states that the supports of the observed variables strictly change with the underlying unobserved component. We believe that this assumption is economically relevant and has the advantages of being simple and testable. Our results have important implications in particular for the measurement error, auction and matching literatures.

Once identification has been established, the problem of estimating these models remain. D'Haultfoeuille and Février (2009), studying common value auctions, propose a multistep nonparametric estimation method that is close to our identification proof. In a first step, the bounds $\underline{S}(.)$ and $\overline{S}(.)$ are estimated, using standard frontier estimation methods. Such methods do not necessarily use solely the observed maxima and minima. Quantiles can be used instead (see Cazals et al., 2005). In a second step, the conditional distribution functions $F_{X|X^*}(.|\underline{X}^{-1}(.))$ are estimated using data for which two "sufficiently extreme" values X_1 and X_2 have been observed. Finally, we use the fact that $\underline{X}^{-1}(.)$ satisfies an integral equation, and estimate it through a Landweber-Fridman estimator (see Carrasco et al., 2007).

Appendix: proofs

Proof of Theorem 2.1

We only prove the second part of the theorem, the first part being an immediate corollary.

First step: identification of $\underline{X}_1(0)$, $\underline{X}_1(1)$, $\overline{X}_2 \circ \underline{X}_1^{-1}(w)$.

By definition, $\underline{X}_k(0) \leq X_k \leq \overline{X}_k(1)$. Hence, $\underline{X}_1(0)$ is identified as the smallest value of X_1. We can also identified $\underline{X}_1(1)$, as the smallest value of X_1 given that X_2 takes its highest value $\overline{X}_2(1)$. Using the same reasoning, $\underline{S}_2(w) = \overline{X}_2 \circ \underline{X}_1^{-1}(w)$ is identified. Given a value w for X_1, the highest value of X_2 compatible with the observation of X_1 is indeed $\overline{X}_2 \circ \underline{X}_1^{-1}(w)$.

To prove that the model is identified, it is thus sufficient to prove that the distribution of X_k ($k \geq 3$) conditional on $W_1 = \underline{X}_1(X^*)$ and the distribution of W_1 are identified. Indeed, for all $w \in [\underline{X}_1(0), \underline{X}_1(1)]$, $f_{X_k|W_1}(x|w) = f_{X_k|X^*}(x|\underline{X}_1^{-1}(w))$ and $F_{W_1}(w) = \underline{X}_1^{-1}(w)$. If both functions are identified, we can then recover $f_{X_k|X^*}(.|.)$.

Second step: identification of the distribution of $X_k|W_1$

For all $\eta > 0$ and $w \in [\underline{X}_1(0), \underline{X}_1(1)]$, let $\underline{w}_\eta = \max(w - \eta, \underline{X}_1(0))$ and $\overline{w}_\eta = \min(w + \eta, \underline{X}_1(1))$. We also define the set $A_\eta(w)$ by

$$A_\eta(w) = \left[\underline{w}_\eta; \overline{w}_\eta\right] \times \left[\underline{S}_2(\underline{w}_\eta); \underline{S}_2(\overline{w}_\eta)\right].$$

When $(X_1, X_2) \in A_\eta(w)$, W_1 belongs to the interval $[\underline{w}_\eta,\ \overline{w}_\eta]$. Indeed, when $W_1 < \underline{w}_\eta$, $X_2 < \underline{S}_2(\underline{w}_\eta)$ and when $W_1 > \overline{w}_\eta$, $X_1 > \overline{w}_\eta$.

For all $\delta > 0$, by continuity of $w \mapsto F_{X_k|W_1}(x|w)$, there exists $\eta > 0$ such that $|F_{X_k|W_1}(x|u) - F_{X_k|W_1}(x|w)| < \delta$ for all u such that $|u - w| < \eta$. Hence,

$$\begin{aligned}
& \left|F_{X_k|(X_1,X_2)\in A_\eta(w)}(x) - F_{X_k|W_1}(x|w)\right| \\
= & \left|\int_{\underline{w}_\eta}^{\overline{w}_\eta} \left(F_{X_k|W_1}(x|u) - F_{X_k|W_1}(x|w)\right) f_{W_1|(X_1,X_2)\in A_\eta(w)}(u)du\right| \\
\leq & \int_{\underline{w}_\eta}^{\overline{w}_\eta} \left|F_{X_k|W_1}(x|u) - F_{X_k|W_1}(x|w)\right| f_{W_1|(X_1,X_2)\in A_\eta(w)}(u)du \\
< & \ \delta,
\end{aligned}$$

where the second line stems from the independence between X_1 and (X_2, X_3) conditional on W_1. Hence, for all $w \in [\underline{X}_1(0), \underline{X}_1(1)]$ and all $x \in [\underline{X}_k \circ \underline{X}_1^{-1}(w), \overline{X}_k \circ \underline{X}_1^{-1}(w)]$,

$$\lim_{\eta \to 0} F_{X_k|(X_1,X_2)\in A_\eta(w)}(x) = F_{X_k|W_1}(x|w).$$

As a consequence, the distribution of X_k conditional on W_1 is identified.

Third step: identification of f_{W_1}

Define

$$q_\eta(w) = P(X_1 \in [\underline{w}_\eta, \overline{w}_\eta]|W = w)P(X_2 \in [\underline{S}_2(\underline{w}_\eta); \underline{S}_2(\overline{w}_\eta)]|W = w).$$

$q_\eta(.)$ is identified by the previous step. Moreover, by conditional independence between X_1 and X_2,

$$P((X_1, X_2) \in A_\eta) = \int_{\underline{w}_\eta}^{\overline{w}_\eta} q_\eta(w) f_{W_1}(w)dw.$$

Let us denote

$$f_{W_1,\eta}(w) = \frac{P((X_1, X_2) \in A_\eta)}{\int_{\underline{w}_\eta}^{\overline{w}_\eta} q_\eta(u)du}.$$

$f_{W_1,\eta}(w)$ is identified, so the result follows if we prove that $\lim_{\eta\to 0} f_{W_1,\eta}(w) = f_{W_1}(w)$.

Because $\underline{X}_1'(.)$ is continuous, so is $f_{W_1}(.)$ on $[\underline{X}_1(0), \underline{X}_1(1)]$. Thus, for all $\delta > 0$, there exists η such that $|u - w| < \eta$ implies that $|f_{W_1}(u) - f_{W_1}(w)| < \delta$. Hence,

$$\begin{aligned}
|f_{W_1,\eta}(w) - f_{W_1}(w)| &= \left| \int_{\underline{w}_\eta}^{\overline{w}_\eta} \frac{q_\eta(u)}{\int_{\underline{w}_\eta}^{\overline{w}_\eta} q_\eta(v)dv} \left(f_{W_1}(u) - f_{W_1}(w)\right) du \right| \\
&\leq \int_{\underline{w}_\eta}^{\overline{w}_\eta} \frac{q_\eta(u)}{\int_{\underline{w}_\eta}^{\overline{w}_\eta} q_\eta(v)dv} \left|f_{W_1}(u) - f_{W_1}(w)\right| du \\
&< \delta \int_{\underline{w}_\eta}^{\overline{w}_\eta} \frac{q_\eta(u)}{\int_{\underline{w}_\eta}^{\overline{w}_\eta} q_\eta(v)dv} du \\
&< \delta.
\end{aligned}$$

The result follows.

Proof of Theorem 4.1

Using $\widetilde{W} = \overline{X}(X^*)$, we shall show that $f_{X|\widetilde{W}}$ and $f_{\widetilde{W}}$ are identified (because of the symmetry, we omit the subscript k on variables whenever unnecessary here). We suppose, without loss of generality, that $K = 3$. Consider (x_1, x_2, x_3) belonging to the support of X and such that $\max(x_1, \overline{X}(0)) \leq x_2 \leq x_3$. Because $\widetilde{W} \geq x_3$, we have

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \int_{x_3}^{\overline{X}(1)} f_{X|\widetilde{W}}(x_1|w) f_{X|\widetilde{W}}(x_2|w) f_{X|\widetilde{W}}(x_3|w) f_{\widetilde{W}}(w)dw$$

For $k \in \{1,2,3\}$, let $\partial_{kr} f_{X_1,X_2,X_3}$ (resp. $\partial_{kl} f_{X_1,X_2,X_3}$) denote the right (resp. left) derivative of f_{X_1,X_2,X_3} with respect to x_k. We have

$$\begin{aligned}\partial_{3l} f_{X_1,X_2,X_3}(x_1,x_2,x_3) = &- f_{X|\widetilde{W}}(x_1|x_3) f_{X|\widetilde{W}}(x_2|x_3) f_{X|\widetilde{W}}(x_3|x_3) f_{\widetilde{W}}(x_3) \\ &+ \int_{x_3}^{\overline{X}(1)} f_{X|\widetilde{W}}(x_1|w) f_{X|\widetilde{W}}(x_2|w) \frac{\partial f_{X|\widetilde{W}}}{\partial x}(x_3|w) f_{\widetilde{W}}(w) dw\end{aligned}$$

Similarly, by taking the derivative in x_2, we find

$$\partial_{2r} f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \int_{x_3}^{\overline{X}(1)} f_{X|\widetilde{W}}(x_1|w) \frac{\partial f_{X|\widetilde{W}}}{\partial x}(x_2|w) f_{X|\widetilde{W}}(x_3|w) dw$$

Hence, if $\underline{X} \leq x \leq w$, we get

$$\partial_{2r} f_{X_1,X_2,X_3}(x,w,w) - \partial_{3l} f_{X_1,X_2,X_3}(x,w,w) = f_{X|\widetilde{W}}(x|w) f^2_{X|\widetilde{W}}(w|w) f_{\widetilde{W}}(w) \tag{5.1}$$

Suppose that $m = 1$, so that $f_{X|\widetilde{W}}(w|w) > 0$. Then

$$f_{X|\widetilde{W}}(x|w) = \frac{\partial_{2r} f_{X_1,X_2,X_3}(x,w,w) - \partial_{3l} f_{X_1,X_2,X_3}(x,w,w)}{\int_{\underline{X}}^{w} \left[\partial_{3r} f_{X_1,X_2,X_3}(y,w,w) - \partial_{2l} f_{X_1,X_2,X_3}(y,w,w)\right] dy}.$$

Because the density f_{X_1,X_2,X_3} and its derivative can be recovered from the data, the right-hand side is identified, and so is $f_{X|\widetilde{W}}(x|w)$.
By (5.1), this implies that $f_{\widetilde{W}}(w)$ is identified by

$$f_{\widetilde{W}}(w) = \frac{\partial_{3r} f_{X_1,X_2,X_3}(x,w,w) - \partial_{2l} f_{X_1,X_2,X_3}(x,w,w)}{f^2_{X|\widetilde{W}}(w|w) f_{X|\widetilde{W}}(x|w)}.$$

Hence, the whole model is identified if $m = 1$.

When $m > 1$, some algebra show that the following equation holds:

$$\frac{\partial f_{X_1,X_2,X_3}}{\partial x_{2r}^{m} \partial x_{3l}^{m-1}}(x,w,w) - \frac{\partial f_{X_1,X_2,X_3}}{\partial x_{2r}^{m-1} \partial x_{3l}^{m}}(x,w,w) = f_{X|\widetilde{W}}(x|w) \left(\frac{\partial f_{X|\widetilde{W}}}{\partial x^{m-1}}\right)^2 (w|w) f_{\widetilde{W}}(w)$$

Hence, reasoning as previously, we also obtain identification in this case.

Proof of Theorem 4.2

It suffices to exhibit a counter-example. Suppose that $X_k = U + \eta_k$ for $k \in \{1,...,K\}$, with $(U, \eta_1, ..., \eta_K)$ mutually independent and $(\eta_1,...,\eta_K)$ identically distributed. We further assume that U has the density function $f_U(x) = (1 - \cos(x))/(\pi x^2)$ and η_k have the density function $f_\eta(x) = f_U(x/2K)/2K$. Given our normalization, this model corresponds to $X_k = F_U^{-1}(X^*) + \eta_k$, with $X^* = F_U(U)$ and F_U being the cumulative distribution function of U. Thus, $f_{X_k|X^*}(x|u) = f_\eta(x - F_U^{-1}(u))$.

Now, let us consider the density function $h_U(x) = f_U(x/2)/6 + 4f_U(2x)/3$. We now show that the model where U has density $h_U(.)$ leads to the same distribution of $(X_1, ..., X_K)$ that the model in which U has density $f_U(.)$. We prove this by showing that the corrresponding characteristic functions coincide. The result follows since the characteristic function uniquely defines the distribution of a random variable.

First, note that the characteristic functions corresponding to f_U, h_U and f_η are respectively $\Psi_U(t) = (1 - |t|)^+$ (where $x^+ = \max(x, 0)$), $\widetilde{\Psi}_U(t) = \frac{1}{3}\Psi_U(2t) + \frac{2}{3}\Psi_U(t/2)$ and $\Psi_\eta(t) = (1 - 2K|t|)^+$. Hence, the characteristic function of $(X_1, ..., X_K)$ satisfies

$$\begin{aligned}\Psi_{X_1,...,X_K}(t_1, ..., t_K) &= \Psi_U\left(\sum_{k=1}^K t_k\right)\prod_{k=1}^K \Psi_\eta(t_k) \\ &= \left(1 - \left|\sum_{k=1}^K t_k\right|\right)^+ \prod_{k=1}^K (1 - 2K|t_k|)^+ .\end{aligned}$$

Now, if $\left|\sum_{k=1}^K t_k\right| \geq 1/2$, $|t_k| \geq 1/2K$ is satisfied for at least one $k \in \{1, ..., K\}$. Hence,

$$\Psi_{X_1,...,X_K}(t_1, ..., t_K) = \Psi_U\left(\sum_{k=1}^K t_k\right)\prod_{k=1}^K \Psi_\eta(t_k) = 0 = \widetilde{\Psi}_U\left(\sum_{k=1}^K t_k\right)\prod_{k=1}^K \Psi_\eta(t_k).$$

Moreover, $\widetilde{\Psi}$ and Ψ_U coincide on $[-1/2, 1/2]$. Indeed, if $\left|\sum_{k=1}^K t_k\right| \leq 1/2$,

$$\begin{aligned}\Psi_U\left(\sum_{k=1}^K t_k\right) &= \left(1 - \left|\sum_{k=1}^K t_k\right|\right)^+ \\ &= \left(1 - \left|\sum_{k=1}^K t_k\right|\right) \\ &= \frac{1}{3}\left(1 - 2\left|\sum_{k=1}^K t_k\right|\right) + \frac{2}{3}\left(1 - \frac{1}{2}\left|\sum_{k=1}^K t_k\right|\right) \\ &= \frac{1}{3}\left(1 - 2\left|\sum_{k=1}^K t_k\right|\right)^+ + \frac{2}{3}\left(1 - \frac{1}{2}\left|\sum_{k=1}^K t_k\right|\right)^+ \\ &= \widetilde{\Psi}_U\left(\sum_{k=1}^K t_k\right)\end{aligned}$$

Hence, for all $(t_1, ..., t_K)$,

$$\Psi_X(t_1, ..., t_K) = \Psi_U\left(\sum_{k=1}^{K} t_k\right)\prod_{k=1}^{K}\Psi_\eta(t_k) = \widetilde{\Psi}_U\left(\sum_{k=1}^{K} t_k\right)\prod_{k=1}^{K}\Psi_\eta(t_k).$$

Going back to densities, this shows that a model where $f_{X_k|X^*}(x|u) = f_\eta(x - H_U^{-1}(u))$, where H_U is the cumulative distribution function corresponding to h_U, leads to the same distribution of $(X_1, ..., X_K)$ as when $f_{X_k|X^*}(x|u) = f_\eta(x - F_U^{-1}(u))$. Because $f_\eta(x - H_U^{-1}(u)) \neq f_\eta(x - F_U^{-1}(u))$, the model is not identified.

References

Becker, G. (1973), 'A theory of marriage: Part i', *Journal of Political Economy* **81**, 813–846.

Bellamn, R. E. and Roth, R. S. (1984), *The Laplace Transform*, World Scientific Publishing Co Pte Ltd.

Bound, J. and Krueger, A. B. (1991), 'The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?', *Journal of Labor Economics* **9**, 1–24.

Carrasco, M., Florens, J. and Renault, E. (2007), Linear inverse problems in structural econometrics, *in* J. Heckman and E. Leamer, eds, 'Handbook of Econometrics', Vol. 6, North Holland.

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006), *Measurement error in nonlinear models: a modern perspective*, Chapman & Hall.

Cazals, C., Florens, J. and Simar, L. (2005), 'Nonparametric frontier estimation: a robust approach', *Journal of Econometrics* **106**, 1–25.

D'Haultfœuille, X. (2009), 'On the completeness condition in nonparametric instrumental regression', *Econometric Theory, forthcoming* .

D'Haultfoeuille, X. and Février, P. (2009), The nonparametric econometrics of common value models. Working Paper.

Embrechts, P., Klueppelberg, C. and Mikosch, T. (1997), *Modelling Extremal Events for insurance and finance*, Springer-Verlag.

Farrell, M. J. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society A* **120**, 253–281.

Février, P., Roos, W. and Visser, M. (2005), 'The revenue effect of the buyer's option in multi-unit ascending auctions', *Journal of Economics and Management Strategy* **14**, 813–847.

Hausman, J., Newey, W. K., Ichimura, H. and Powell, J. (1991), 'Measurement errors in polynomial regression models', *Journal of Econometrics* **50**, 273–295.

Hu, Y. (2008), 'Identification and estimation of nonlinear models with misclassification error using instrumental variables: a general solution', *Journal of Econometrics* **144**, 27–61.

Hu, Y., McAdams, D. and Shum, M. (2009), Nonparametric identification of auction models with nonseparable unobserved heterogeneity. Working Paper 553, Department of Economics, Johns Hopkins University.

Hu, Y. and Schennach, S. (2008), 'Identification and estimation of nonclassical errors-in-variable models with continuous distributions using instruments', *Econometrica* **76**, 195–216.

Hu, Y. and Shum, M. (2009*a*), Estimating first-price auctions with unknown number of bidders: a misclassification approach. Working Paper 541, Department of Economics, Johns Hopkins University.

Hu, Y. and Shum, M. (2009*b*), Nonparametric identification of dynamic models with unobserved state variables. Working Paper 543, Department of Economics, Johns Hopkins University.

Krasnokutskaya, E. (2009), 'Identification and estimation in highway procurement auctions under unobserved auction heterogeneity', *Review of Economic Studies* **Forthcoming**.

Li, T. (2002), 'Robust and consistent estimation of nonlinear errors-in-variables models', *Journal of Econometrics* **110**, 1–26.

Newey, W. K. (2001), 'Flexible simulated moment estimation of nonlinear errors-in-variables models', *Review of Economics and Statistics* **83**, 616–627.

Newey, W. and Powell, J. (2003), 'Instrumental variable estimation of nonparametric models', *Econometrica* **71**, 1565–1578.

Riley, J. and Samuelson, W. (1981), 'Optimal auctions', *The American Economic Review* **71**, 381–392.

Rodgers, W. L., Brown, C. and Duncan, G. J. (1993), 'Errors in survey reports of earnings, hours worked, and hourly wages', *Journal of the American Statistical Association* **88**, 1208–1218.

Schennach, S. (2004), 'Estimation of nonlinear models with measurement error', *Econometrica* **72**, 33–75.

Schennach, S. (2007), 'Instrumental variables estimation of nonlinear errors-in-variables models', *Econometrica* **75**, 201–239.

Shimer, R. and Smith, L. (2000), 'Assortative matching and search', *Econometrica* **68**, 343–369.