



WORKING PAPERS

N° 1218

May 2021

“Non Parametric Classes for Identification in Random
Coefficients Models when Regressors have Limited
Variation”

Christophe Gaillac and Eric Gautier

NONPARAMETRIC CLASSES FOR IDENTIFICATION IN RANDOM COEFFICIENTS MODELS WHEN REGRESSORS HAVE LIMITED VARIATION

CHRISTOPHE GAILLAC⁽¹⁾ AND ERIC GAUTIER⁽¹⁾

ABSTRACT. This paper studies point identification of the distribution of the coefficients in some random coefficients models with exogenous regressors when their support is a proper subset, possibly discrete but countable. We exhibit trade-offs between restrictions on the distribution of the random coefficients and the support of the regressors. We consider linear models including those with nonlinear transforms of a baseline regressor, with an infinite number of regressors and deconvolution, the binary choice model, and panel data models such as single-index panel data models and an extension of the Kotlarski lemma.

1. INTRODUCTION

Random coefficients models are used to incorporate multiple sources of unobserved heterogeneity in modelling various economic behaviors. In these models, the researcher can be interested in more than the average of the vector of random coefficients and even might want to recover its distribution [see, *e.g.*, 6, 24, 19, 32, 40, 14]. Imposing parametric assumptions on the law of the random coefficients is a widely used approach in the analysis of random coefficients models [see, *e.g.*, 5, 57] but can seriously drive the results [see 28, 9]. Economic theory rarely motivates such restrictions. When the coefficients are random with a law in a nonparametric class and independent from the regressors, identification often requires the regressors to have a support which is the whole space. However, in applications regressors may only have limited variation.

The main contribution of this paper is to provide identification of the distribution of the coefficients in some random coefficients models with regressors independent from the random coefficients, namely exogenous ones, when their support is a proper subset, possibly discrete but countable. Estimation results along those lines for the linear model are available in [21]. We provide a general identification strategy which is then used to study linear models including those with nonlinear transforms of a baseline regressor, with an infinite number of regressors and deconvolution, the binary choice model, and panel data models such as single-index panel data models and an extension of the Kotlarski lemma. We show in these models that the support of the regressors can be discrete if we maintain integrability

⁽¹⁾ TOULOUSE SCHOOL OF ECONOMICS, UNIVERSITÉ TOULOUSE CAPITOLE, 1 ESPLANADE DE L'UNIVERSITÉ, 31000 TOULOUSE, FRANCE.

E-mail addresses: christophe.gaillac@tse-fr.eu, eric.gautier@tse-fr.eu.

Date: This version: May 20, 2021.

Keywords: Identification, Random Coefficients, Quasi-analyticity, Deconvolution.

AMS 2010 Subject Classification: Primary 62P20 ; secondary 42A99, 62G07, 62G08.

The authors acknowledge financial support from the grants ERC POEMH 337665 and ANR-17-EURE-0010. Christophe Gaillac thanks CREST/ENSAE where this research was partly conducted.

assumptions on the distribution of the coefficients. Using several examples, we illustrate the trade-off between the variation the regressors and the restrictions on the nonparametric class of distributions of random coefficients we consider.

The paper is organized as follows. Section 2 introduces the notations and the nonparametric identification strategy. In Section 3, we provide the main identification results for the various linear random coefficients models. Section 4 provides results on the random coefficients binary model and Section 5 on some panel data models with random coefficients. The appendix provides complements on the tools used for identification and the proofs.

2. PRELIMINARIES

Bold letters are used for vectors, capital letters for indeterminates of polynomials or random variables/vectors. For a real number r , \mathbf{r}_c is the vector, the dimension of which will be clear from the text, where each entry is r . The notations \mathbb{N} and \mathbb{N}_0 are used for the positive and nonnegative integers, for $p \in \mathbb{N}$ $[p] = \{1, \dots, p\}$, $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$, $\mathbb{C}(X)$ denotes the set of rational functions, and $\mathbb{1}\{\cdot\}$ the indicator function. $\mathbb{K}[Z_1, \dots, Z_p]$ is the ring of polynomials of p variables with coefficients in the ring \mathbb{K} . Let $S \subseteq \mathbb{R}^p$, $C(S)$ and $C^\infty(S)$ be the continuous and infinitely differentiable functions at every point in S with values in \mathbb{C} . For $\mathcal{S} \subseteq \mathbb{C}^p$, $\mathcal{A}(\mathcal{S})$ and $\mathcal{H}^\infty(\mathcal{S})$ are the analytic and bounded analytic functions on \mathcal{S} . Quasi-analytic classes of functions on \mathbb{R}^p are vector spaces of complex valued functions in $C^\infty(\mathbb{R}^p)$ characterized by the knowledge of their derivatives at $\mathbf{0}_c$. Unlike real analytic functions, the Taylor series of such functions do not need to converge. The notation $|\cdot|_q$ for $q \in [1, \infty]$ stands for the ℓ_q norm of a vector with components in \mathbb{C} while $L^q(S, \mu)$ for $q \in [1, \infty]$ are the q integrable functions on S with respect to the measure μ and the norms $\|\cdot\|_{L^q(S, \mu)}$. When μ is the Lebesgue measure we drop it from the notation. S^p is the unit sphere in \mathbb{R}^{p+1} . $(\mathbf{e}_j)_{j=1}^p$ is the canonical basis of \mathbb{R}^p . For $\boldsymbol{\beta} \in \mathbb{C}^p$, $k \in \mathbb{N}_0$, and $\mathbf{m} \in \mathbb{N}_0^p$, denote by $\mathbf{m}! = \prod_{j=1}^p \mathbf{m}_j!$, $|\mathbf{m}| = \sum_{j=1}^p \mathbf{m}_j$, $\boldsymbol{\beta}^{\mathbf{m}} = \prod_{j=1}^p \boldsymbol{\beta}_j^{\mathbf{m}_j}$, and $|\boldsymbol{\beta}|^{\mathbf{m}} = \prod_{j=1}^p |\boldsymbol{\beta}_j|^{\mathbf{m}_j}$. $GL(\mathbb{R}^p)$ are the invertible $p \times p$ matrices with real coefficients.

$\mathfrak{M}_c(\Omega)$, $\mathfrak{M}(\Omega)$, and $\mathfrak{M}_1(\Omega)$ are the sets of complex, nonnegative, and probability measures on a Borel measurable set Ω . When $\Omega \subset \mathbb{R}^p$ is closed, these are measures over \mathbb{R}^p with support in Ω . For $\mu \in \mathfrak{M}_c(\mathbb{R}^p)$ and $\mathbf{m} \in \mathbb{N}_0^p$, $|\mu|$ is its total variation, $s_\mu(\mathbf{m}) = \int_{\mathbb{R}^p} \mathbf{x}^{\mathbf{m}} d\mu(\mathbf{x})$ the moments, and by $s_{|\cdot|, |\mu|}(\mathbf{m}) = \int_{\mathbb{R}^p} |\mathbf{x}^{\mathbf{m}}| d|\mu|(\mathbf{x})$ the absolute moments. Define

$$\mathfrak{M}_c^*(\mathbb{R}^p) = \left\{ \mu \in \mathfrak{M}_c(\mathbb{R}^p) : \forall \mathbf{m} \in \mathbb{N}^p, \sum_{j=1}^p s_{|\cdot|, |\mu|}(2\mathbf{m}\mathbf{e}_j) < \infty \right\},$$

$$\mathfrak{M}^*(\mathbb{R}^p) = \mathfrak{M}_c^*(\mathbb{R}^p) \cap \mathfrak{M}(\mathbb{R}^p), \quad \mathfrak{M}_1^*(\mathbb{R}^p) = \mathfrak{M}^*(\mathbb{R}^p) \cap \mathfrak{M}_1(\mathbb{R}^p).$$

Let \mathcal{M} be $\mathfrak{M}_c^*(\Omega)$ or $\mathfrak{M}^*(\Omega)$. A measure μ is said to be determinate in \mathcal{M} if $\mu \in \mathcal{M}$ and if μ is uniquely determined in \mathcal{M} by $\{s_\mu(\mathbf{m})\}_{\mathbf{m} \in \mathbb{N}_0^p}$.

The Fourier transform of $\mu \in \mathfrak{M}_c(\mathbb{R}^d)$ (resp. f in $L^q(\mathbb{R}^d)$ for $q = 1, 2$) is $\mathcal{F}[\mu] : \mathbf{x} \mapsto \int_{\mathbb{R}^d} e^{i\mathbf{b}^\top \mathbf{x}} d\mu(\mathbf{b})$ (resp. $\mathcal{F}[f]$). For a random vector \mathbf{X} , $\mathbb{P}_{\mathbf{X}}$ is its law, $F_{\mathbf{X}}$ its CDF, $f_{\mathbf{X}}$ its density, $\varphi_{\mathbf{X}} = \mathcal{F}[\mathbb{P}_{\mathbf{X}}]$, and $\mathbb{S}_{\mathbf{X}}$ its support. $\mathbb{E}_{\mathbb{P}}[\cdot]$ is the expectation under \mathbb{P} , \otimes the product of measures, and $\mathbb{P}_{Y|\mathbf{X}}(\cdot|\mathbf{x})$ for $\mathbf{x} \in \mathbb{S}_{\mathbf{X}}$ the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$.

This paper considers models of the form

$$(1) \quad \mathbf{Y} = \mathbf{v}(\mathbf{X}, \boldsymbol{\Gamma}), \quad \mathbf{X} \perp\!\!\!\perp \boldsymbol{\Gamma},$$

where \mathbf{v} is a known measurable vector valued function, \mathbf{Y} is the vector of outcomes, $\mathbf{\Gamma} \in \Gamma$ and $\mathbf{X} \in \mathcal{X}$ are vectors of unobserved and observed factors, and Γ and \mathcal{X} are the Euclidian space or the sphere. Everything in this paper holds if we impose independence given \mathbf{Z} , where \mathbf{Z} is a random vector from which we could have observations simultaneously with those of \mathbf{Y} and \mathbf{X} or which could be identifiable from a model for \mathbf{X} obtained by a control function approach. The independence means that \mathbf{X} is exogenous, it yields

$$\forall \mathbf{x} \in \mathbb{S}_{\mathbf{X}}, \mathbb{P}_{\mathbf{Y}|\mathbf{X}}(\cdot|\mathbf{x}) = \mathbb{P}_{\mathbf{v}(\mathbf{x},\mathbf{\Gamma})}.$$

In order to study identification, we proceed as if $(\mathbb{P}_{\mathbf{Y}|\mathbf{X}=\mathbf{x}})_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}}$ were known and denote the true law by $\mathbb{P}_{\mathbf{\Gamma}}^*$. We maintain restrictions \mathcal{R} on the primitives $\mathbb{P}_{\mathbf{X}} \otimes \mathbb{P}_{\mathbf{\Gamma}}$ which involve two conditions: (i) $V \subseteq \mathbb{S}_{\mathbf{X}}$ and (ii) $\mathbb{P}_{\mathbf{\Gamma}} \in \mathcal{P}$, for well chosen sets V and \mathcal{P} . We use the notation $(\mathbb{S}_{\mathbf{X}}, \mathbb{P}_{\mathbf{\Gamma}}) \in \mathcal{R}$. Random coefficients models involve inner products of subvectors of $\mathbf{\Gamma}$ and \mathbf{X} , hence indices with random coefficients, and possibly additional idiosyncratic errors. The linear random coefficients model can be written as

$$(2) \quad Y = \alpha + \boldsymbol{\beta}^\top \mathbf{X}, \quad \mathbf{\Gamma} = (\alpha, \boldsymbol{\beta}^\top)^\top \perp\!\!\!\perp \mathbf{X}.$$

It is a baseline component of more complex random coefficient models. For example, [23] shows how identification of a discrete choice model with random coefficients and a special regressor is amenable to identification of (2) while without special regressor requires a treatment like in Section 4. Obtaining identification of $\mathbb{P}_{\mathbf{\Gamma}}^*$ in these non separable models with multiple unobservables without restricting \mathcal{P} usually requires that $\mathbb{S}_{\mathbf{X}} = \mathcal{X}$. For model (2) this can be shown by the Cramer-Wold Theorem. $\mathbb{S}_{\mathbf{X}} = \mathcal{X}$ is too demanding for a dataset. However, identification can hold when $\mathbb{S}_{\mathbf{X}}$ is a proper subset but \mathcal{P} is a restricted nonparametric class. Because random coefficient models involve indices, we can study identification when $\mathbb{S}_{\mathbf{X}}$ is replaced by a convenient invertible affine transformation. So when the results in this paper depend on the order of the regressors, the order is irrelevant. Indeed, for all $\underline{\mathbf{x}} \in \mathbb{R}^p$ and $M \in GL(\mathbb{R}^p)$,

$$(3) \quad \alpha + \boldsymbol{\beta}^\top \mathbf{X} = \alpha + \boldsymbol{\beta}^\top \underline{\mathbf{x}} + \boldsymbol{\beta}^\top M^{-1}M(\mathbf{X} - \underline{\mathbf{x}})$$

and there is a one to one relation between $\mathbb{P}_{\alpha + \boldsymbol{\beta}^\top \underline{\mathbf{x}}, (M^{-1})^\top \boldsymbol{\beta}}$ and $\mathbb{P}_{\alpha, \boldsymbol{\beta}}$.

Let $\mathcal{T}_{\mathbb{S}_{\mathbf{X}}}^v$ map $\mathbb{P}_{\mathbf{\Gamma}}$ to the collection of direct image measures $(\mathbb{P}_{\mathbf{v}(\mathbf{x},\mathbf{\Gamma})})_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}}$. $\mathbb{P}_{\mathbf{\Gamma}}^*$ is identified under \mathcal{R} if $\mathcal{T}_{\mathbb{S}_{\mathbf{X}}}^v$ is injective. The proofs that we give often rely on the diagram

$$(4) \quad \begin{array}{ccc} \mathbb{P}_{\mathbf{\Gamma}} \in \mathcal{P} & \xrightarrow{\mathcal{T}_{\mathbb{S}_{\mathbf{X}}}^v} & (\mathbb{P}_{\mathbf{v}(\mathbf{x},\mathbf{\Gamma})})_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}} \\ \downarrow \mathcal{G} \text{ (injective)} & & \downarrow \Pi \\ \mathcal{G}[\mathbb{P}_{\mathbf{\Gamma}}] \in \mathfrak{F}(S) & \xrightarrow{\text{restriction to } U} & \mathcal{G}[\mathbb{P}_{\mathbf{\Gamma}}]_U \end{array}$$

The choice of Π depends on \mathbf{v} . By restricting \mathcal{P} , $\mathcal{G}[\mathbb{P}_{\mathbf{\Gamma}}]$ belongs to a class of functions $\mathfrak{F}(S)$ which also contains all differences between functions in $\mathcal{G}[\mathbb{P}_{\mathbf{\Gamma}}]$ (e.g., $\mathfrak{F}(S)$ could be a vector space), \mathcal{G} is injective, and U is a set of uniqueness of $\mathfrak{F}(S)$.

Definition 1. $U \subseteq S$ is a set of uniqueness for a vector space of functions $\mathfrak{F}(S)$ on S if every function of $\mathfrak{F}(S)$ which is zero on U is identically zero on S .

3. THE LINEAR RANDOM COEFFICIENTS MODEL

The model is (2). It is a natural specification to account for heterogenous effects. It can be viewed as more general than the quantile regression. Indeed, when the conditional quantiles are strictly increasing, the quantile regression defines the same data generating process as a linear random coefficients model where the coefficients are functions of a scalar uniform distribution. The unobserved scalar uniform variable is a ranking variable. (2) allows for the coefficients to be a function of a multidimensional vector, possibly infinite. [23] allows for extensive generalizations of (2) including ones involving nonlinear transforms of a baseline regressor. This paper considers different specifications. In Section 3.2, (2) is an approximation of the nonlinear model

$$(5) \quad \boxed{Y = g(\mathbf{X}, \Theta), \Theta \perp\!\!\!\perp \mathbf{X}},$$

where g is unknown, Θ has arbitrary, possibly infinite, dimension. Indeed, when $\mathbb{S}_{\mathbf{X}}$ is compact and almost surely in $\theta \in \mathbb{S}_{\Theta}$, $\mathbf{x} \mapsto g(\mathbf{x}, \theta) \in L^2(\mathbb{S}_{\mathbf{X}})$, $(f_j)_{j \in \mathbb{Z}}$ is a Riesz basis and $f_0 = 1$, then

$$(6) \quad g(\mathbf{x}, \theta) = \gamma_0(\theta) + \sum_{j \in \mathbb{Z} \setminus \{0\}} \gamma_j(\theta) f_j(\mathbf{x}).$$

Recall that a Riesz basis of a separable Hilbert space is the image of an orthonormal basis by a bounded invertible operator and that the coefficients $\gamma_j(\theta)$ are the inner products of $\mathbf{x} \mapsto g(\mathbf{x}, \theta)$ with a uniquely defined biorthogonal system (see sections 1.7 and 1.8 in [60]). An approximation of (5) is thus

$$(7) \quad Y = \gamma_0(\Theta) + \sum_{j \in J} \gamma_j(\Theta) f_j(\mathbf{X}), \Theta \perp\!\!\!\perp \mathbf{X},$$

where $J \subset \mathbb{Z} \setminus \{0\}$ is finite. Equation (7) can be used for extrapolation of the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$ when the functions f_j are analytic (see also [20]). Hence, it is possible to extrapolate not only conditional expectation functions but also any conditional quantiles even under a nonparametric specification with multiple unobservables entering in a non-additively separable way. In (5) or its approximation (7), the researcher can also be interested in the law of the random coefficients to eventually obtain the law of the elasticity $\partial_x g(x, \Theta) x / g(x, \Theta)$ or of the marginal effects $\partial_x g(x, \Theta)$ assuming they exist and x belongs to the interior of $\mathbb{S}_{\mathbf{X}}$.

Remark 1. A variation of (7) consists in taking $J = \mathbb{Z} \setminus \{0\}$ but there exists $j_0 \in \mathbb{N}$ such that $\gamma_j(\Theta)$ for $|j| > j_0$ are deterministic and $\sum_{j \in \mathbb{Z}} \mathbb{E} [\gamma_j(\Theta)^2] < \infty$. $\mathbb{E} [\gamma_j(\theta)]$ are identified because $\mathbb{E} [\gamma_j(\theta)] = \mathbb{E} [Y g_j(\mathbf{X}) / f_{\mathbf{X}}(\mathbf{X})]$, where $(g_j)_{j \in \mathbb{Z}}$ is the biorthogonal system associated to $(f_j)_{j \in \mathbb{Z}}$. So this model is amenable to (7) with an a priori unknown set J . However, by taking j_0 to be the smallest k such that

$$j_0 = \operatorname{argmin}_{k \in \mathbb{N}} \inf_{\mathbb{P}_{\mathbf{r}}} \int_{\mathbb{R}} \mathbb{E} \left[\left(\mathbb{P}(Y \leq y | \mathbf{X}) - \mathbb{P} \left(\sum_{|j| \leq k} \Gamma_j f_j(\mathbf{X}) \leq y \right) \right)^2 \right] dy.$$

Finally, we also consider in this section infinite dimensional linear random coefficients models. There, the random coefficients belongs to $\ell_2(\mathbb{N}_0)$ a.s. It is denoted by $\{\Gamma_m\}$ and is

such that $\Gamma_0 = \alpha$ and, for all $m \in \mathbb{N}$, $\Gamma_m = \beta_m$. The regressors $\{X_m\}$ belong to $\ell_2(\mathbb{N})$ a.s. and the model can be written

$$(8) \quad Y = \alpha + \sum_{m=1}^{\infty} \beta_m X_m, \quad \{\Gamma_m\} \perp\!\!\!\perp \{X_m\}.$$

It can be related to the random coefficients functional linear regression (see [59])

$$Y = \int_{-1}^1 \pi(s) X(s) ds + U,$$

where a.s. $X \in L^2(-1, 1)$, $\pi = \sum_{m=0}^{\infty} \Pi_m \nu_m \in L^2(-1, 1)$ and $X_m := \int_{-1}^1 \nu_k(s) X(s) ds$ belong respectively to $\ell_2(\mathbb{N}_0)$ and $\ell_2(\mathbb{N})$, and $(\nu_m)_{m \in \mathbb{N}_0} \in L^2(-1, 1)^{\mathbb{N}_0}$. Similar expressions are used for model (8).

In this section, we consider different identifying restrictions which we denote by $\mathcal{R}_{L,j}$, for $j \in [7]$ and are defined below. The main theorem of this section is the following.

Theorem 1. \mathbb{P}_{Γ}^* in (2) is identified under either of $\mathcal{R}_{L,j}$ for $j \in \{1, 2, 3, 6, 7\}$. $\mathbb{P}_{\{\Gamma_m\}}^*$ in (8) is identified under either of $\mathcal{R}_{L,j}$ for $j \in \{4, 5\}$. Moreover, \mathbb{P}_{Γ}^* in (2) is not identified if $\mathcal{R}_{L,3}$ (ii) holds but not $\mathcal{R}_{L,3}$ (i).

The proofs rely on diagram (4) with $\mathcal{G} = \mathcal{F}$ (sometimes extended to an open set of \mathbb{C}^{p+1}), $U = \{(t, t\mathbf{x}), (t, \mathbf{x}) \in \mathbb{R} \times \mathbb{S}_{\mathbf{X}}\}$, and

$$\Pi : \left(\mathbb{P}_{(1, \mathbf{x}^\top) \Gamma} \right)_{\mathbf{x} \in \mathbb{S}_{\mathbf{X}}} \mapsto \left((t, \mathbf{x}) \in \mathbb{R} \times \mathbb{S}_{\mathbf{X}} \mapsto \int_{\mathbb{R}^{p+1}} e^{it(1, \mathbf{x}^\top) \gamma} d\mathbb{P}_{\Gamma}(\gamma) \right),$$

with the relevant modifications when dealing with sequences and model (8). We denote by $\mathcal{P}(\Omega)$ (resp. $\mathcal{P}_c(\Omega)$) the set of (resp. complex) measures with support included in Ω which are determinate in $\mathfrak{M}^*(\Omega)$ (resp. $\mathfrak{M}_c^*(\Omega)$). Appendix A.1.2 reviews criteria for determinacy and classical examples are given in Remark 2. The vector space spanned by $\mathcal{F}[\mathcal{P}(\Omega)]$ (resp. $\mathcal{F}[\mathcal{P}_c(\Omega)]$) is a quasi-analytic class of functions on \mathbb{R}^p . It is known that sets with a nonempty interior and, when $p = 1$, sets which contain a bounded sequence of distinct points are sets of uniqueness. For analytic classes this is the Weierstrass theorem [see, e.g., Theorem 15.11 in 47]) and for quasi-analytic classes this follows by the arguments in the proof of Lemma 4.8 in [1] for the quasi-analytic class under consideration. We now give details on the restrictions.

3.1. Regressors with finite support and independence of the marginals of \mathbb{P}_{Γ} , deconvolution. This is a challenging situation in terms of limited variation of the regressors. We present restrictions in the spirit of [2]. Let $\Omega_0 \subseteq \mathbb{R}$ and, for all $k \in [p]$, $\Omega_k \subseteq \mathbb{R}$ be given closed sets. These sets account for possible prior information on the support of the marginals of \mathbb{P}_{Γ} . When such an information is not available, we take the set to be \mathbb{R} . We consider either of the two following restrictions.

Restriction 1. $\mathcal{R}_{L,1}$ (i) $\{\mathbf{0}_c, \mathbf{x}(1), \dots, \mathbf{x}(p)\} \subseteq \mathbb{S}_{\mathbf{X}}$ and $(\mathbf{x}(1), \dots, \mathbf{x}(p))$ is an upper triangular matrix with nonzero diagonal elements;

$$\mathcal{R}_{L,1} \text{ (ii)} \quad \mathcal{P} = \{\mathbb{P}_{\Gamma} = \mathbb{P}_{\alpha} \otimes \bigotimes_{k=1}^p \mathbb{P}_{\beta_k} : \forall k \in [p], \mathbb{P}_{\beta_k} \in \mathcal{P}(\Omega_k)\}.$$

Restriction 2. $\mathcal{R}_{L,2}$ (i) $\mathcal{R}_{L,1}$ (i) holds ;

$$\mathcal{R}_{L,2} \text{ (ii)} \quad \mathcal{P} = \{\mathbb{P}_{\Gamma} = \mathbb{P}_{\alpha} \otimes \bigotimes_{k=1}^p \mathbb{P}_{\beta_k} : \mathbb{P}_{\alpha} \in \mathcal{P}(\Omega_0), \forall k \in [p-1], \mathbb{P}_{\beta_k} \in \mathcal{P}(\Omega_k)\}.$$

The problem of deconvolution with two samples, one of the error and one of the sum of the signal and the error, fits model (2) with $p = 1$ and $\mathbb{S}_X = \{0, 1\}$ is a particular case (see [23]). $\mathcal{R}_{L,2}$ (ii) is in the spirit of the assumption used in [10]. Under $\mathcal{R}_{L,1}$ (ii), φ_α can have zeros on an open set at the expense of a stronger assumption on \mathbb{P}_β . There are classical examples of characteristic functions with compact support (e.g., $t \mapsto (1 - |t|^r)\mathbb{1}\{|t| \leq 1\}$ for $0 < r \leq 1$). Remark 1 in [36] relies on analyticity. [41] considers the estimation of compactly supported densities of β which implies $\mathcal{R}_{L,1}$ (ii) and φ_β is analytic.

Now on, we do not assume mutual independence of the random coefficients. This is important if Γ is of the form $\Gamma(\Theta)$, where Θ is a deep heterogeneity parameter Θ as in (7).

3.2. \mathbb{S}_X is not in the zeros of a nonzero polynomial and nonlinear model. Let $\Omega \subseteq \mathbb{R}^{p+1}$ be a closed set and the researcher knows that $\mathbb{S}_\Gamma \subseteq C$.

Restriction 3. $\mathcal{R}_{L,3}$ (i) \mathbb{S}_X is not a subset of the zeros of a nonzero element of $\mathbb{R}[Z_1, \dots, Z_p]$; $\mathcal{R}_{L,3}$ (ii) \mathcal{P} is the set of measures which are determinate in $\mathfrak{M}^*(\Omega)$.

Lemma 2 in [40] considers the case where $\Omega = \mathbb{R}^{p+1}$ and assumes \mathbb{S}_X contains an open ball which implies $\mathcal{R}_{L,3}$ (i) (see also [3]). $\mathcal{R}_{L,3}$ (ii) is much weaker than assuming that \mathbb{S}_Γ is compact as in [3]. By the usual properties of the Fourier transform, the vector space spanned by $\mathcal{F}[\mathcal{P}]$ is a quasi-analytic class. Because $\mathcal{A}(\mathbb{R}^{p+1})$ is a small subset of $C^\infty(\mathbb{R}^{p+1})$ (see Appendix A.1.4), it is important in the analysis of the paper to use larger classes and allow for Fourier transforms to not be analytic. The next examples satisfy $\mathcal{R}_{L,3}$ (i) (i.e. are sets of uniqueness of $\mathbb{R}_d[Z_1, \dots, Z_p]$).

Example 1. (i) $p = 1$ and $|\mathbb{S}_X| = \infty$;

(ii) $p \geq 2$ and $U_p \subseteq \mathbb{S}_X$, where U_p is defined recursively via U_1 , such that $|U_1| = \infty$ and, for all $j = 2, \dots, p$, $U_j = \bigcup_{\mathbf{u} \in U_{j-1}} \{(\mathbf{u}^\top, v)^\top, v \in V_j(\mathbf{u})\}$, where $|V_j(\mathbf{u})| = \infty$.

This is a set of uniqueness because $P \in \mathbb{R}_d[Z_1, \dots, Z_p]$ can be written as $P(Z_1, \dots, Z_p) = \sum_{k=0}^d Q_k(Z_1, \dots, Z_{p-1})Z_p^k$, where $Q_k \in \mathbb{R}_d[Z_1, \dots, Z_{p-1}]$. Examples of sets U_p are $\prod_{k=1}^p V_k \subseteq \mathbb{S}_X$, where, for all $k \in [p]$, $|V_k| = \infty$, the infinite fan $\{\mathbf{x} \in \mathbb{R}^2 : \exists n \in \mathbb{N}, \mathbf{x}_2 = n\mathbf{x}_1\}$, and infinite staircase $\{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}_2 = \lceil \mathbf{x}_1 \rceil\}$ (see [7]).

A consequence of the second statement in Theorem 1 is that \mathbb{P}_Γ^* in (2) when $\mathbf{X} = (X \ X^2)^\top$ under $\mathcal{R}_{L,3}$ (ii) holds, even if $\mathbb{S}_X = \mathbb{R}$. Indeed, \mathbb{S}_X is included in the set of zeros of $Q(Z_1, Z_2) = Z_1^2 - Z_2$. Example 2 shows that we can handle other nonlinear transformations of a baseline variable which can have discrete support. It applies to model (7).

Example 2. $\{(f_1(\mathbf{u}), \dots, f_p(\mathbf{u}))^\top, \mathbf{u} \in U_q\} \subseteq \mathbb{S}_X$, where $q \in \mathbb{N}$, $\mathcal{B}(\mathcal{S})$ is a vector space of functions on $\mathcal{S} \subseteq \mathbb{C}^q$ which is an algebra with respect to multiplication, the set of functions $\{f_j, j \in [p]\}$ of $\mathcal{B}(\mathcal{S})$ is such that if $P(f_1, \dots, f_p) = 0$, with $P \in \mathbb{R}[Z_1, \dots, Z_p]$, then $P = 0$, and U_q a set of uniqueness of $\mathcal{B}(\mathcal{S})$ (see Section A.1.3 for classes of analytic and quasi-analytic functions).

Because $\mathcal{B}(\mathcal{S})$ is an algebra with respect to multiplication, $P(f_1, \dots, f_p) \in \mathcal{B}(\mathcal{S})$. Because U_q is a set of uniqueness of $\mathcal{B}(\mathcal{S})$, if $P(f_1, \dots, f_p)(\mathbf{x}) = 0$ for all $\mathbf{x} \in U_q$ then $P(f_1, \dots, f_p) = 0$, so $P = 0$. Hence, $\mathcal{R}_{L,3}$ (i) holds. Here, p can be large and U_q discrete. The last condition in Example 2 means that $\{f_j, j \in [p]\}$ is algebraically independent over \mathbb{R} . Clearly, algebraic independence over k implies algebraic independence over \mathbb{R} if \mathbb{R} is a subfield of k . We now give useful examples to apply Example 2.

If $f_1(z) = z_1, \dots, f_{p-1}(z) = z_q, f_p$ is a real or complex analytic function of z_1, \dots, z_{p-1} and $\{f_j, j \in [p]\}$ is a set of algebraically independent over \mathbb{C} , f_p is said to be transcendental in \mathbb{C} . If f is a transcendental function in \mathbb{C} , then so is $1/f$ and its inverse on a domain where it is invertible. If f is meromorphic in \mathbb{C} then it is a transcendental function in \mathbb{C} iff it is not a rational function (e.g., $\exp(z)/z, \sin(z)/(z-1)^2, \cosh(z), \Gamma(z), \zeta(z)$). Classical examples of transcendental functions include $z^\pi, \exp(z), z^z, z^{1/z}$, and $\log(z)$.

If $q \leq p-1$ and $f_1(x) = x_1, \dots, f_q(x) = x_q$, then $\{f_j, j \in [p]\}$ is an algebraically independent set over \mathbb{C} iff $\{f_j, j \in \{q+1, \dots, p\}\}$ is algebraically independent over $\mathbb{C}(Z_1, \dots, Z_q)$. For example, $\{f_j, j \in [p-1]\}$ are algebraically independent over $\mathbb{C}(Z)$ if:

- (i) $f_j(z) = e^{\lambda_j z}$, where $(\lambda_j)_{j=1}^p$ are such that, if $(b_j)_{j \in J} \in \mathbb{N}_0^p$ is such that $\sum_{j=1}^p b_j \lambda_j = 0$, then, for all $j = 1, \dots, p, b_j = 0$ (see proof of Proposition 1).
- (ii) $f_j = e^{\varphi_j}$ and $\varphi_j \in \mathcal{A}_0(\mathcal{S})$, where $\mathcal{A}_0(\mathcal{S}) \subseteq \mathcal{A}(\mathcal{S})$ and $\mathcal{S} \subseteq \mathbb{C}$ is a simply connected open subset which does not contain 0 but a set of uniqueness of $\mathcal{A}_0(\mathcal{S})$, such that, for $j = 2, \dots, p, \lim_{x \rightarrow \infty} |\varphi_j(x)/\varphi_{j-1}(x)| = \infty$ and $\lim_{n \rightarrow \infty} |\varphi_1(x)/\log(x)| = \infty$. Thus, $\{\Psi, e^{\varphi_1 \circ \Psi}, \dots, e^{\varphi_p \circ \Psi}\}$ is algebraically independent over \mathbb{C} on a connected open subset $\mathcal{S}_1 \subseteq \mathbb{C}$ if $\Psi : \mathcal{S}_1 \rightarrow \mathcal{S}$ is analytic.

A way to achieve the condition in (i) above is to rely on a transcendental number r over \mathbb{Z} (e.g., e or π). It means that, for all $P \in \mathbb{Z}[Z], P(r) = 0$ implies that $P = 0$. The following result shows that we can work with the specification (7) with functions from a Riesz basis which are algebraically independent.

Proposition 1. *Let $T \in \mathbb{R}_+, \lambda_j = j + 1/(5r^{|j|})$ and $f_j(z) = e^{i\pi\lambda_j z/T}$, for all $j \in \mathbb{Z} \setminus \{0\}$ and $r \in (1, \infty)$ is transcendental over \mathbb{Z} , and $f_0 = 1$. We have:*

- (P6.1) $\{f_j : j \in \mathbb{Z} \setminus \{0\}\}$ is an algebraically independent family of functions over $\mathbb{C}(X)$;
- (P6.2) $(f_j)_{j \in \mathbb{Z}}$ is a Riesz basis of $L^2(-T, T)$.

(P6.1) implies $\{f_j : j \in \mathbb{Z} \setminus \{0\}\}$ is an algebraically independent family of functions over \mathbb{R} which is the condition in Example 2. Here $q = 1$ and U_q should be such that $U_q \subseteq [-T, T]$.

Consider model (8). $(\{e_{j,m}\})_{j \in \mathbb{N}}$ is now the canonical basis of $\ell_2(\mathbb{N})$. Let G be a vector subspace of $\ell_2(\mathbb{N})$ and μ a probability measure on $\ell_2(\mathbb{N})$, then we denote by $\Pi_{G*}\mu$ the projection of μ onto G .

Restriction 4. *There exists $n_0 \in \mathbb{N}$ and $\{\underline{x}_m\} \in \ell_2(\mathbb{N})$ such that, for all $n \geq n_0$,*

- $\mathcal{R}_{L,4}$ (i) $\{(x_1, \dots, x_n) : \{x_m\} \in \mathbb{S}_{\{X_m\}}, \forall j > n, x_j = \underline{x}_j\}$ is not a subset of the zeros of a nonzero element of $\mathbb{R}[Z_1, \dots, Z_n]$;
- $\mathcal{R}_{L,4}$ (ii) \mathcal{P} is the set of measures whose projections onto $G_{n+1} := \text{Span}(\{e_{1,m}\}, \dots, \{e_{n+1,m}\})$ are determinate in $\mathfrak{M}^*(\mathbb{R}^{n+1})$.

$\mathcal{R}_{L,4}$ (i) allows discrete support of the regressors. Like in Section 3.1, less restrictive conditions on $\mathbb{S}_{\{X_m\}}$ can be obtained assuming independence between marginals of $\mathbb{P}_{\{\Gamma_m\}}$. An alternative restriction is:

- Restriction 5.** $\mathcal{R}_{L,5}$ (i) *There exists a Gaussian measure μ on $\ell_2(\mathbb{N})$ such that $\{\{u_m\} \in \mathbb{R}^{\mathbb{N}_0} : \forall n \in \mathbb{N}, u_n = u_0 x_n, \{x_m\} \in \mathbb{S}_{\{X_m\}}\}$ has positive μ -measure;*
- $\mathcal{R}_{L,5}$ (ii) \mathcal{P} is the set of measures such that, for all $n \in \mathbb{N}$ and $\{f_m\} \in \ell_2(\mathbb{N}) \setminus \{\{0_m\}\}$, the projections onto $G_{n+1} := \text{Span}(\{e_{1,m}\}, \dots, \{e_{n,m}\}, \{f_m\})$ are determinate in $\mathfrak{M}^*(\mathbb{R}^{n+1})$.

3.3. Without restrictions on α . $\mathcal{R}_{L,3}$ (ii) places restrictions on \mathbb{P}_α which we entirely remove. This is important to model income or wealth and allow α to have, for example, a Pareto distribution which does not even belong to $\mathfrak{M}^*(\mathbb{R})$. We can replace the role of α by β_k for $k \in [p]$ or a combination of the coefficients if, starting from (3), we form, for example,

$$\frac{Y}{\mathbf{X}_k - \mathbf{x}_k} = \beta_k + \frac{\alpha}{\mathbf{X}_k - \mathbf{x}_k} + \sum_{j \neq k} \beta_j \frac{\mathbf{X}_j - \mathbf{x}_j}{\mathbf{X}_k - \mathbf{x}_k}$$

and there exists $\mathbf{x}_k \notin \mathbb{S}_{\mathbf{X}_k}$. If the support of a subvector of dimension k of \mathbf{X} is \mathbb{R}^k then we can similarly handle situations where there are no restrictions on the joint distribution of the corresponding coefficients and α . In this section, $\Omega \subseteq \mathbb{R}^p$ is a closed set and it can be known that $\mathbb{S}_\beta \subseteq \Omega$.

A key idea is to rely on the partial Fourier transform. For example, $\mathcal{F}[\mathbb{P}_\Gamma]$ can be analyzed as the collection of functions $\mathcal{F}[\mathbb{P}_\Gamma](t, \cdot)$ for $t \in \mathbb{R}$ and

$$\forall t \in \mathbb{R}, \mathcal{F}[\mathbb{P}_\Gamma](t, \star) = \mathcal{F}[\mathbb{P}_{\beta,t}](\star), \text{ where } \mathbb{P}_{\beta,t} = \int_{\mathbb{R}} e^{ita} d\mathbb{P}_\Gamma(a, \cdot).$$

Because the vector space spanned by $\mathcal{F}[\mathcal{P}_c(\Omega)]$ is a quasi-analytic class of functions on \mathbb{R}^p and by the continuity of the Fourier transform at 0, \mathbb{P}_Γ^* in (2) is identified under $\mathcal{R}_{L,6}$ below.

Restriction 6. $\mathcal{R}_{L,6}$ (i) $\mathbb{S}_{\mathbf{X}}$ has a nonempty interior or, when $p = 1$, $\mathbb{S}_{\mathbf{X}}$ contains a bounded sequence of distinct points;

$\mathcal{R}_{L,6}$ (ii) \mathcal{P} restricts \mathbb{P}_Γ so that $\mathbb{P}_{\beta,t} \in \mathcal{P}_c(\Omega)$ for all $t \neq 0$.

$\mathcal{R}_{L,6}$ (i) implies that $t\mathbb{S}_{\mathbf{X}}$ is a set of uniqueness of quasi-analytic classes for all $t \neq 0$.

In order to have quantitative statements on $\mathbb{S}_{\mathbf{X}}$ yielding sets of uniqueness and explicit conditions so that $\mathcal{R}_{L,6}$ (ii) holds, we restrict our attention to the case where there are sets of points $V_k \subseteq \mathbb{R}$ for $k \in [p]$ such that $\prod_{k=1}^p V_k \subset \mathbb{S}_{\mathbf{X}}$. The following restriction involves pairs (C_k, \mathcal{P}_k) for $k \in [p]$. \mathcal{P}_k are classes of probabilities of the form

$$\mathcal{P}_k = \left\{ \mathbb{P}_{\beta_k} \in \mathfrak{M}_1^*(\Omega_k) : \forall h \in \mathcal{H}_k, \mathbb{E}[h(\beta_k)] \leq M_k(h) \right\},$$

where $\Omega_k \subseteq \mathbb{R}$ is a closed set, which can be taken not equal to \mathbb{R} if the researcher knows $\mathbb{S}_{\beta_k} \subseteq \Omega_k$, and \mathcal{H}_k can consist of one or a sequence of nonnegative measurable functions. The parameters of these classes are Ω_k , \mathcal{H}_k , and M_k . The proof relies on the related class

$$\mathcal{P}_{c,k} = \left\{ \mu \in \mathfrak{M}_c^*(\Omega_k) : \forall h \in \mathcal{H}_k, \int_{\mathbb{R}} h(z) d|\mu|(z) \leq 2M_k(h) \right\}.$$

The classes C_k are such that $\mathcal{F}[\mathcal{P}_{c,k}] \subseteq C_k \subseteq C^\infty(\mathbb{R})$.

Restriction 7. $\mathbb{S}_{\mathbf{X}}$ and \mathcal{P} are such that, for all $k \in [p]$,

$\mathcal{R}_{L,7}$ (i) tV_k is a set of uniqueness of C_k for all $t \neq 0$;

$\mathcal{R}_{L,7}$ (ii) $\mathbb{P}_{\beta_k} \in \mathcal{P}_k$.

Let us give examples of pairs $(V_k, \mathcal{P}_k)_{k=1}^p$ and of corresponding sets of uniqueness. Some examples use log-convex sequences $(M_m)_{m \in \mathbb{N}_0}$ which are sequences of nonnegative numbers such that, for all $m \in \mathbb{N}$, $M_m^2 \leq M_{m-1}M_{m+1}$ and we also assume $M_0 = 1$. Define also the iterated logarithms by $\log^{*0}(t) = t$ and for $j \geq 1$ by $\log^{*j}(t) = \log(\log^{*(j-1)}(t))$ provided t is large enough for the quantity to be defined and \exp^{*n} is the reciprocal function.

(E.1) For $k \in [p]$, $\mathbb{E} [e^{r|\beta_k|}] \leq m(r)$ for all $r \in \mathbb{R}_+$, where $\lim_{r \rightarrow \infty} m(r) = \infty$ and V_k satisfies

$$(9) \quad \forall t > 0, \exists \alpha > 1 : \overline{\lim}_{r \rightarrow \infty} \frac{\log(\alpha)}{\log(m(\alpha tr))} |V_k \cap ((-r, r) \setminus \{0\})| > 1.$$

(E.2) For $k \in [p]$, there exists $(b, c) \in \mathbb{R}_+^2$ such that $\mathbb{E} [|\beta_k|^m] \leq cb^m m!$ for all $m \in \mathbb{N}_0$ and V_k satisfies

$$(10) \quad \overline{\lim}_{r \rightarrow \infty} \frac{\log(|V_k \cap (-r, r)|)}{r} = \infty.$$

(E.3) For $k \in [p]$, $\mathbb{S}_{\beta_k} \subseteq \mathbb{R}_+$, there exist $(b, c) \in \mathbb{R}_+^2$ and a log-convex sequence $\{M_m\}$ such that $M_0 = 1$, $\mathbb{E} [|\beta_k|^m] \leq cb^m M_m$ for all $m \in \mathbb{N}_0$,

$$(11) \quad \sum_{m \in \mathbb{N}} \frac{1}{M_m^{1/(2m)}} = \infty,$$

and V_k contains a bounded sequence of distinct points;

(E.4) For $k \in [p]$, there exist $(b, c) \in \mathbb{R}_+^2$ and a log-convex sequence $\{M_m\}$ such that $M_0 = 1$, $\mathbb{E} [|\beta_k|^m] \leq cb^m M_m$ for all $m \in \mathbb{N}_0$, and either

(E.4a) $M_m = \nu(m)m!$, where ν satisfies, for all $n \in \mathbb{N}$, $\nu(m) = 1$ for $0 \leq m < \exp^{*n}(1)$ and else $\nu(m) = (\log(m) \log^{*2}(m) \cdot \log^{*n}(m))^m$, and V_k satisfies

$$\overline{\lim}_{r \rightarrow \infty} \frac{\log^{*n+1}(|V_k \cap (-r, r)|)}{r} = \infty;$$

(E.4b) or $\{M_m\}$ satisfies

$$(12) \quad \sum_{m \in \mathbb{N}} \frac{1}{M_{2m}^{1/(2m)}} = \infty$$

and V_k contains a bounded sequence of distinct points;

In (E.1), a sufficient condition for (9) when $\mathbb{S}_{\beta_k} \subseteq [-\rho, \rho]$, hence $m(r) = e^{\rho r}$, is

$$\overline{\lim}_{r \rightarrow \infty} \frac{1}{r} |V_k \cap ((-r, r) \setminus \{0\})| = \infty.$$

In (E.2), C_k consists of the functions f such that there exists $\rho > 0$ such that f can be extended uniquely to $\mathcal{H}^\infty(\{z \in \mathbb{C} : |\operatorname{Im}(z)| < \rho\})$ (see Theorem 19.9 in [47]).

In (E.3), and (E.4), C_k are quasi-analytic classes which are detailed in Appendix A.1. In (E.3), using Proposition A.1 (2) with $\{M_m\} = \{m!\}$, C_k is an analytic class and the set of uniqueness is given in (2) in Section 1 of [31]. In particular, Proposition A.1 shows that $\mathcal{F}[\mathcal{P}_{c,k}] \subseteq C_k$. Based on Theorem 4b in (author?), (year?), Example 2 in Appendix A.1 gives more general forms of ν in the definition of $\{M_m\}$ in (E.4a) and associated sets V_k . The relation between condition (12) (resp. (11)) and the determinacy of measures in $\mathfrak{M}^*(\mathbb{R})$ (resp. $\mathfrak{M}^*(\mathbb{R}_+)$) is explicited in Section A.1.2.

Remark 2. *The Student's t with $0 < \nu < \infty$ degrees of freedom, the generalized gamma $GG(a, b, p)$ for $0 < a < 1/2$ of density $ab^p x^{ap-1} \exp(-bx^a)/\Gamma(p)$ on \mathbb{R}^+ , any positive power of the lognormal, the law of N^{2n+1} for all $n \in \mathbb{N}$, $|N|^r$ for $r > 4$, X^m for all $m \in \mathbb{N} \setminus \{1, 2\}$, Y^r for all $|r| > 2$, where N is a Gaussian, X a Laplace, gamma or logistic, and Y an inverse Gaussian random variables, respectively do not satisfy these conditions. However, $|N|^r$ for all $0 < r \leq 4$, X^m for $m = 1, 2$, Y^r for all $-2 \leq r \leq 2$, $GG(a, b, p)$ for all $a \geq 1/2$ (thus the χ^2 with any degrees of freedom) satisfy these conditions, hence are determinate*

in the space of measures with the appropriate support restriction (i.e., $\mathfrak{M}^*(\mathbb{R})$ or $\mathfrak{M}^*(\mathbb{R}_+)$) [see 54, 44, 35, for more examples].

4. THE RANDOM COEFFICIENTS BINARY CHOICE MODEL

The random coefficients binary choice model takes the form

$$(13) \quad Y = \mathbb{1}\{\alpha + \boldsymbol{\beta}^\top \mathbf{X} \geq 0\}, \quad (\alpha, \boldsymbol{\beta}) \perp \mathbf{X}.$$

The binary variable Y is 1 when an individual chooses a good, treatment, or an action. In the context of treatment effects, $Y = 1$ when an individual chooses treatment. [58] shows the monotonicity of [34] is equivalent to a selection equation with an additively separable latent index with a single unobservable. Hence monotonicity can be related to rank invariance (see [13]) similarly to the fact that, in Section 3, a linear random coefficients models where the random coefficients are functions of a scalar unobservable can be related to the quantile regression. It is well known that monotonicity is sometimes too strong and it is exemplified in [34]. [29] call (13) the benchmark nonseparable, nonmonotonic model of treatment choice. Because the scale of the index is not identified, various normalizations can be used. A simple one used in consists in assuming one coefficient is 1. This requires that in the original scale the coefficient has a (strict) sign a.s. Based on this and under sufficient variation of the corresponding (special) regressor [23] show that identification corresponds to identification of the model of Section 3. It was mentioned that this idea applies to all sorts of models involving as constitutive element random coefficients indices lying in certain rectangles (e.g., choice models with multiple alternatives or entry games). [23] also considers the case where the special regressor and the remaining ones have limited variation. The scaling of [24] in a preliminary version [23] was removed and overlooked as not important. However, assuming that one coefficient is 1 is unnecessary and too restrictive because it still imposes some form of monotonicity. Some elements can be found in [22] in the context of selection equations and missing data in sample surveys.

Assuming that $\mathbb{P}(|(\alpha, \boldsymbol{\beta}^\top)^\top|_2 = 0) = 0$, (13) can be equivalently written as

$$(14) \quad Y = \mathbb{1}\{\boldsymbol{\Gamma}^\top \mathbf{S} \geq 0\}, \quad \boldsymbol{\Gamma} \perp \mathbf{S},$$

where $\boldsymbol{\Gamma} = (\alpha, \boldsymbol{\beta}^\top)^\top / |(\alpha, \boldsymbol{\beta}^\top)^\top|_2$ and $\mathbf{S} = (1, \mathbf{X}^\top)^\top / |(1, \mathbf{X}^\top)^\top|_2$. Clearly $|(1, \mathbf{X}^\top)^\top|_2 \geq 1$ and the support of \mathbf{S} is a closed subset of the hemisphere $H^+ = \{\mathbf{s} \in \mathbb{S}^p : \mathbf{s}_1 \geq 0\}$. We consider identification of the density $f_{\boldsymbol{\Gamma}}^*$ of $\mathbb{P}_{\boldsymbol{\Gamma}}^*$ with respect to σ , which is the surface measure on \mathbb{S}^p . In this section, we consider the following restriction of the class \mathcal{P}

$$\mathcal{P}_{BC} = \{\mathbb{P}_{\boldsymbol{\Gamma}} \in \mathfrak{M}_1(\mathbb{S}^p) : d\mathbb{P}_{\boldsymbol{\Gamma}} = f_{\boldsymbol{\Gamma}} d\sigma, f_{\boldsymbol{\Gamma}}(\mathbf{u})f_{\boldsymbol{\Gamma}}(-\mathbf{u}) = 0 \text{ for a.e. } \mathbf{u} \in \mathbb{S}^p\}.$$

It is shown in [25] that $\mathbb{P}_{\boldsymbol{\Gamma}}^*$ is identified under the restriction $\mathbb{S}_{\mathbf{X}} = \mathbb{R}^p$ and $\mathcal{P} = \mathcal{P}_{BC}$. It is assumed in [24] that the support of $\boldsymbol{\Gamma}$ lies in an (unknown) hemisphere, namely, that there exists \mathbf{n} in \mathbb{S}^p such that $\mathbb{P}(\mathbf{n}^\top \boldsymbol{\Gamma} \geq 0) = 1$. This assumption first appeared in [33] and is just a sufficient but unnecessary assumption, similar to assuming that one coefficient is 1. Indeed, if $\mathbf{n} \in H^+$, then we have $\mathbb{P}(Y = 1 | \mathbf{S} = \mathbf{n}) = 1$, else $\mathbb{P}(Y = 1 | \mathbf{S} = -\mathbf{n}) = 0$. This means that there exist limits of values of the regressors such that in the limit everyone chooses $Y = 1$ or in the limit everyone chooses $Y = 0$. It is stronger than $f_{\boldsymbol{\Gamma}}(\mathbf{u})f_{\boldsymbol{\Gamma}}(-\mathbf{u}) = 0$ for a.e. \mathbf{u} in \mathbb{S}^p which does not imply “unselected samples”.

Consider now the case where the support of \mathbf{X} is a proper subset of \mathbb{R}^p or equivalently the support of \mathbf{S} is a proper subset of H^+ . The hemispherical transform is defined, for

$\mathbf{s} \in \mathbb{S}^p$ and $f \in L^1(\mathbb{S}^p)$, by $\mathcal{T}f(\mathbf{s}) = \int_{\mathbb{S}^p} \mathbb{1}\{\mathbf{u}^\top \mathbf{s} \geq 0\} f(\mathbf{u}) d\sigma(\mathbf{u}) - 1/2$. Using the restrictions \mathcal{P}_{BC} , we have

$$(15) \quad \forall \mathbf{s} \in \mathbb{S}_{\mathcal{S}}, \quad \mathcal{T}f_{\Gamma}(\mathbf{s}) + \frac{1}{2} = \mathbb{E}[Y|\mathcal{S} = \mathbf{s}].$$

V -quasi-analytic classes of functions on \mathbb{S}^p are vector spaces of functions on \mathbb{S}^p characterized by $(\Delta^m f(\mathbf{x}))_{m \in \mathbb{N}_0}$ for all $\mathbf{x} \in V$.

In this section, we consider two different identifying restrictions which we denote by $\mathcal{R}_{BC,j}$ for $j \in [2]$. The main theorem of this section is the following

Theorem 2. \mathbb{P}_{Γ}^* in (14) is identified under either of $\mathcal{R}_{BC,j}$ for $j \in \{1, 2\}$.

Similarly to Section 3, the proofs rely on the diagram (4) with $\mathcal{G} = \mathcal{T}$, $U = \{\mathbf{s} = (1, \mathbf{x}^\top) / |(1, \mathbf{x}^\top)|_2, \mathbf{x} \in \mathbb{S}_{\mathcal{X}}\}$, and

$$\Pi : \left(\mathbb{P}_{\mathbb{1}\{\Gamma^\top \mathbf{s} \geq 0\}} \right)_{\mathbf{s} \in \mathbb{S}_{\mathcal{S}}} \mapsto \left(\mathbf{s} \in \mathbb{S}^p \mapsto \int_{\mathbb{S}^p} \mathbb{1}\{\mathbf{u}^\top \mathbf{s} \geq 0\} f(\mathbf{u}) d\sigma(\mathbf{u}) - \frac{1}{2} \right).$$

In the rest of this section, we consider $\mathcal{P} \subseteq \mathcal{P}_{BC}$ such that, for given $V \subseteq \mathbb{S}^p$, the vector space spanned by $\mathcal{T}[f_{\Gamma}]$ for all $\mathbb{P}_{\Gamma} \in \mathcal{P}$ is a V -quasi-analytic class of functions on \mathbb{S}^p . We now give details on the restrictions.

For functions on \mathbb{S}^p , the Laplacian Δ has eigenspaces $H_{m,p}$, eigenvalues $\zeta_{m,p} = -m(m + p - 1)$, and $Q_{m,p}f(\cdot) = \int_{\mathbb{S}^p} q_{m,p}(\cdot, \mathbf{y}) f(\mathbf{y}) d\sigma(\mathbf{y})$ is the orthogonal projection of f onto $H^{m,p}$ for all $m \in \mathbb{N}_0$.

Restriction 8. Let $0 < \epsilon < 1$.

$\mathcal{R}_{BC,1}$ (i) $\mathbb{S}_{\mathcal{X}}$ is a set of uniqueness of $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C}^p : |\text{Im}(\mathbf{z})|_2 < \epsilon\})$;

$\mathcal{R}_{BC,1}$ (ii) \mathcal{P} is such that $\overline{\lim}_{m \rightarrow \infty} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)}^{1/m} < 1/(1 + 2\epsilon)$.

$\mathcal{R}_{BC,1}$ (ii) is a sufficient condition for: $\mathbb{E}[Y|\mathbf{X} = \cdot]$ belongs to $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C}^p : |\text{Im}(\mathbf{z})|_2 < \epsilon\})$. Clearly, a set of uniqueness of $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C}^p : |\text{Im}(\mathbf{z})|_2 < \epsilon\})$ is a set of uniqueness of the superset $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C}^p : |\text{Im}(\mathbf{z})|_\infty < \epsilon\})$. Hence, a sufficient condition for $\mathcal{R}_{BC,1}$ (i) is that $U_p \subseteq \mathbb{S}_{\mathcal{X}}$ where for all $j = 2, \dots, p$, $U_j = \bigcup_{\mathbf{u} \in U_{j-1}} \{(\mathbf{u}^\top, v)^\top, v \in V_j(\mathbf{u})\}$, where $V_j(\mathbf{u})$ and U_1 are sets of uniqueness of $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C} : |\text{Im}(\mathbf{z})| < \epsilon\})$. A particular case is a product of sets of uniqueness of $\mathcal{H}^\infty(\{\mathbf{z} \in \mathbb{C} : |\text{Im}(\mathbf{z})| < \epsilon\})$. We give examples in Section A.1.3.

The odd part f^- of $f \in L^q(\mathbb{S}^p)$ is the limit in $L^q(\mathbb{S}^p)$ of $f_n^-(\mathbf{x}) = (f_n(\mathbf{x}) - f_n(-\mathbf{x}))/2$, where $(f_n)_{n \in \mathbb{N}_0} \in (C(\mathbb{S}^p))^{\mathbb{N}_0}$ converges to f in $L^q(\mathbb{S}^p)$. We make use of $C_{\text{odd}}^\infty(\mathbb{S}^p)$, the restriction to odd functions of $C^\infty(\mathbb{S}^p)$, and

$$C(\{M_m\}) = \{f \in C^\infty(\mathbb{S}^p) : \exists c, b \in \mathbb{R}_+ : \forall m \in \mathbb{N}_0, \|\Delta^m f^-\|_{L^1(\mathbb{S}^p)} \leq cb^m M_m\}.$$

Restriction 9. $\{M_m\}$ satisfies (12) and

$\mathcal{R}_{BC,2}$ (i) There exists $U \subseteq \mathbb{S}_{\mathcal{S}}$ a set of uniqueness of harmonic homogenous polynomials of odd degree in $\mathbb{R}[Z_1, \dots, Z_{p+1}]$ and, for all $f \in C_{\text{odd}}^\infty(\mathbb{S}^p) \cap \mathcal{T}[C(\{M_m\})]$, if, for all $\mathbf{u} \in \mathbb{S}_{\mathcal{S}}$, $f(\mathbf{u}) = 0$, then, for all $\mathbf{u} \in U$ and $m \in \mathbb{N}_0$, $\Delta^m f(\mathbf{u}) = 0$;

$\mathcal{R}_{BC,2}$ (ii) \mathcal{P} is such that $f_{\Gamma} \in C(\{M_m\})$.

U is a set of uniqueness of homogeneous polynomials in $\mathbb{R}[Z_1, \dots, Z_{p+1}]$ if there exists $A \in GL(\mathbb{R}^{p+1})$ such that $\{(1, \mathbf{u}^\top)^\top : \mathbf{u} \in \tilde{U}\} \subseteq AU$ and \tilde{U} is a set of uniqueness of $\mathbb{R}[Z_1, \dots, Z_p]$. If $\mathbb{S}_{\mathcal{X}}$ and thus $\mathbb{S}_{\mathcal{S}}$ contains an open set, then clearly $\mathcal{R}_{BC,2}$ (i) is satisfied.

5. PANEL DATA MODELS WITH LINEAR RANDOM COEFFICIENTS MODEL

5.1. **Extension of the Kotlarski lemma.** Consider the equation

$$(16) \quad \mathbf{Y}_t = \delta + \boldsymbol{\epsilon}_t, \quad t \in \{1, 2\},$$

where the vector of unobserved heterogeneity is $\boldsymbol{\Gamma} := (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \delta)$.

Restriction 10. Define the restriction \mathcal{R}_K on \mathcal{P} by $\{\mathbb{P}_{\boldsymbol{\Gamma}} = \mathbb{P}_{\boldsymbol{\epsilon}_1} \otimes \mathbb{P}_{\boldsymbol{\epsilon}_2} \otimes \mathbb{P}_{\delta}, \mathbb{P}_{\boldsymbol{\epsilon}_1} \in \mathcal{P}_1(\Omega_1)\}$, where $\Omega_1 \subseteq \mathbb{R}$ is a closed set and $\mathcal{P}_1(\Omega_1)$ is the set of measures which are determinate in $\mathfrak{M}^*(\Omega_1)$ and such that $\mathbb{E}_{\mathbb{P}_{\boldsymbol{\epsilon}_1}}[\boldsymbol{\epsilon}_1] = 0$

Theorem 3. $\mathbb{P}_{\boldsymbol{\Gamma}}^*$ in (16) is identified under \mathcal{R}_K .

Lemma 1 in [36] assumes all characteristic functions do not vanish and in Remark 1 it is written that this can be extended to the case where all characteristic functions are analytic. Theorem 3 shows that these assumptions are too strong and identification can be achieved when none of the characteristic functions are analytic and $\boldsymbol{\epsilon}_2$ and δ might not have finite first absolute moments. [18] present a similar result under alternative assumptions, but assuming $\varphi_{\boldsymbol{\epsilon}_1}$ is analytic.

5.2. **Linear panel data model where regressors are monomials of a baseline scalar regressor.** Consider the equation

$$(17) \quad \mathbf{Y}_t = \alpha + \sum_{j=1}^T \beta_j \mathbf{X}_t^j + \boldsymbol{\epsilon}_t, \quad t = 1, \dots, T,$$

where \mathbf{X}_t is a scalar regressor and denote by $\boldsymbol{\Gamma} = (\alpha, \boldsymbol{\beta}^\top)^\top$. For each t , we have

$$\mathbb{S}_{\mathbf{X}_t, \dots, \mathbf{X}_t^T} \subseteq \{\mathbf{u} \in \mathbb{R}^p : \mathbf{u}_2 = \mathbf{u}_1^2, \dots, \mathbf{u}_T = \mathbf{u}_1^T\},$$

hence the restriction $\mathcal{R}_{L,3}$ does not hold. We show in this section that the availability of T periods allows nonparametric identification.

Remark 3. Note that (17) could be generalized to $\mathbf{Y}_t = \alpha + P(\mathbf{X}_t) + \boldsymbol{\epsilon}_t$, for all $t = 1, \dots, T$, where $P(\mathbf{X}_t) = \sum_{j=1}^T \beta_j \mathbf{X}_t^{\theta(j)}$ and $\theta \in \mathbb{N}_0^{\mathbb{N}}$ and increasing. However, using our identification strategy would yield to consider the so-called generalized Vandermonde matrices, whose theoretical properties, in particular the ones of their inverse, are not yet well known.

Restriction 11. $\mathcal{R}_{LP,0}$ (i) $\mathbb{P}_{\boldsymbol{\Gamma}} = \bigotimes_{j=1}^T \mathbb{P}_{\boldsymbol{\epsilon}_j} \otimes \mathbb{P}_{\alpha, \boldsymbol{\beta}}$, $\mathbb{P}_{\boldsymbol{\epsilon}_1} \in \mathcal{P}_1(\Omega_1)$, where $\Omega_1 \subseteq \mathbb{R}$ is a closed set and $\mathcal{P}_1(\Omega_1)$ is the set of measures which are determinate in $\mathfrak{M}^*(\Omega_1)$ and such that $\mathbb{E}_{\mathbb{P}_{\boldsymbol{\epsilon}_1}}[\boldsymbol{\epsilon}_1] = 0$;

$\mathcal{R}_{LP,0}$ (ii) $\mathcal{X}_1 = \{\mathbf{x} \in \mathbb{S}_{\mathbf{X}} : \mathbf{x}_1 = \dots = \mathbf{x}_T\} \neq \emptyset$.

Using an extension of Theorem 3 to T periods, see proof of Theorem 4 below, $\mathbb{P}_{\boldsymbol{\Gamma}}^*$ is identified under $\mathcal{R}_{LP,0}$. Note that the restriction $\mathcal{R}_{LP,0}$ (ii) is weaker than assuming that the covariates are centered $\mathbf{0}_c \in \mathbb{S}_{\mathbf{X}}$. The restriction $\mathcal{R}_{LP,0}$ (ii) is also maintained in [14], where they focus on the marginals of $\boldsymbol{\Gamma}$ without imposing our baseline independence assumption but considering only the individuals whose \mathbf{X} belong to \mathcal{X}_1 , which are the stayers.

Restriction 12. \mathcal{R}_{LP} (i) For all $t = 1, \dots, T$, the support of \mathbf{X}_t contains a bounded sequence of distinct points;

\mathcal{R}_{LP} (ii) $\mathbb{P}_{\beta_T} \in \mathcal{P}_T(\Omega_T)$, where $\Omega_T \subseteq \mathbb{R}$ is a closed set and $\mathcal{P}_T(\Omega_T)$ is the set of measures which are determinate in $\mathfrak{M}^*(\Omega_T)$.

Theorem 4. $\mathbb{P}_{\Gamma, \epsilon}^*$ in (17) is identified under $\mathcal{R}_{LP,0}$ and \mathcal{R}_{LP} .

Like Theorem 1 under $\mathcal{R}_{L,6}$ or $\mathcal{R}_{L,7}$, Theorem 4 makes no assumptions on $\mathbb{P}_{\alpha, \beta_1, \dots, \beta_{T-1}}^*$.

5.3. A single-index panel data model with two periods. Consider the equation

$$(18) \quad \mathbf{Y}_t = f(\mathbf{\Gamma}^\top \mathbf{X}_t) + \boldsymbol{\eta}_t, \quad t = 1, 2, \quad f \text{ is increasing.}$$

Restriction 13. Let $0 < \epsilon < 1$. Define the restriction based on both $\mathcal{R}_{BC,2}$ and

- \mathcal{R}_{SI} (i) $\{(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{S}_{\mathbf{X}_1, \mathbf{X}_2} : \mathbf{x}_1 = \mathbf{x}_2\} \neq \emptyset$;
 \mathcal{R}_{SI} (ii) (a) $\mathbb{P}_{\Gamma, \eta} = \mathbb{P}_{\Gamma} \otimes \mathbb{P}_{\eta_1} \otimes \mathbb{P}_{\eta_2}$, $\mathbb{P}_{\eta_1} \in \mathcal{P}_1(\Omega_1)$, where $\Omega_1 \subseteq \mathbb{R}$ is a closed set and $\mathcal{P}_1(\Omega_1)$ is the set of measures which are determinate in $\mathfrak{M}^*(\Omega_1)$, $\mathbb{E}_{\mathbb{P}_{\eta_1}}[\boldsymbol{\eta}_1] = 0$;
 (b) $|\mathbf{\Gamma}|_2 = 1$.

Theorem 5. $\mathbb{P}_{\Gamma, \eta}^*$ in (18) is identified under \mathcal{R}_{SI} .

Theorem 5 relies on both theorems 2 and 3. Restriction \mathcal{R}_{SI} (i) means that there exist “stayers” in the population, for which the value of the covariate stays the same across periods, which is a mild assumption.

REFERENCES

- [1] C. Belisle, J.-C. Massé, and T. Ransford. When is a probability measure determined by infinitely many projections? *Annals of Probability*, 25:767–786, 1997.
- [2] R. Beran and P. Hall. Estimating coefficient distributions in random coefficient regressions. *Annals of Statistics*, 20:1970–1984, 1992.
- [3] R. Beran and W. Millar. Minimum distance estimation in random coefficient regression models. *Annals of Statistics*, 22:1976–1992, 1994.
- [4] C. Berg. *The multidimensional moment problem and semi-groups*, In: Landau, H.J. (ed.) *Moments in Mathematics*. American Mathematical Society, 1992.
- [5] S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, pages 841–890, 1995.
- [6] S. T. Berry and P. A. Haile. Nonparametric identification of multinomial choice demand models with heterogeneous consumers. Technical report, National Bureau of Economic Research, 2009.
- [7] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*. Springer, 1998.
- [8] S. Bochner and A. Taylor. Some theorems on quasi-analyticity for functions of several variables. *American Journal of Mathematics*, 61:303–329, 1939.
- [9] C. Breunig. Varying random coefficient models. *Journal of Econometrics*, 221(2):381–408, 2021.
- [10] M. Carrasco and J.-P. Florens. A spectral method for deconvolving a density. *Econometric Theory*, 27:546–581, 2011.
- [11] F. S. Cater. Differentiable, nowhere analytic functions. *American Mathematical Monthly*, 91:618–624, 1984.
- [12] I. Chalendar and J. R. Partington. Multivariable approximate Carleman-type theorems for complex measures. *Annals of Probability*, 35:384–396, 2007.

- [13] V. Chernozhukov and C. Hansen. An iv model of quantile treatment effets. *Econometrica*, 73:245–261, 2005.
- [14] J. Coopridge, S. Hoderlein, and A. Meister. A panel data estimator for the distribution and quantiles of marginal effects in nonlinear structural models with an application to the demand for junk food. *Available at SSRN 3545485*, 2020.
- [15] J. A. Cuesta-Albertos, R. Fraiman, and T. Ransford. A sharp form of the Cramer-Wold theorem. *Journal of Theoretical Probability*, 20:201–209, 2007.
- [16] M. De Jeu. Determinate multidimensional measures, the extended Carleman theorem and quasi-analytic weights. *Annals of Probability*, 20:1205–1227, 2003.
- [17] M. De Jeu. Subspaces with equal closure. *Constructive Approximation*, 20:93–157, 2004.
- [18] K. Evdokimov and H. White. Some extensions of a lemma of Kotlarski. *Econometric Theory*, 28:925–932, 2012.
- [19] J. Fox. A note on nonparametric identification of distributions of random coefficients in multinomial choice models. *NBER working paper series*, 2017.
- [20] C. Gaillac and E. Gautier. Estimates for the SVD of the truncated fourier transform on $L^2(\cosh(b\cdot))$ and stable analytic continuation. 2019. Preprint [arXiv:1905.11338v3](https://arxiv.org/abs/1905.11338v3).
- [21] C. Gaillac and E. Gautier. Adaptive estimation in the linear random coefficients model when regressors have limited variation. *Forthcoming, Bernoulli*, 2021.
- [22] E. Gautier. Relaxing monotonicity in endogenous selection models and application to surveys. In *Advances in Contemporary Statistics and Econometrics*, 2021.
- [23] E. Gautier and S. Hoderlein. A triangular treatment effect model with random coefficients in the selection equation. 2011,2015. Preprint [arXiv:1109.0362v4](https://arxiv.org/abs/1109.0362v4).
- [24] E. Gautier and Y. Kitamura. Nonparametric estimation in random coefficients binary choice models. *Econometrica*, 81:581–607, 2013.
- [25] E. Gautier and E. Le Pennec. Adaptive estimation in the nonparametric random coefficients binary choice model by needlet thresholding. *Electronic Journal of Statistics*, 12:277–320, 2018.
- [26] G. H. Hardy. On Stieltjes’ “problème des moments”. *Messenger of Math.*, 46:175–182, 1917.
- [27] G. H. Hardy. On Stieltjes’ “problème des moments”. *Messenger of Math.*, 47:84–91, 1918.
- [28] J. Heckman and B. Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52:271–320, 1984.
- [29] J. Heckman and E. Vytlacil. Structural equations, treatment effects, and econometric policy evaluation1. *Econometrica*, 73:669–738, 2005.
- [30] C. C. Heyde. Some remarks on the moment problem. *The Quarterly Journal of Mathematics*, 14:91–96, 1963.
- [31] I. I. Hirschman. On the distribution of the zeros of functions belonging to certain quasi-analytic classes. *American Journal of Mathematics*, 72:396–406, 1950.
- [32] S. Hoderlein, H. Holzmann, and A. Meister. The triangular model with random coefficients. *Journal of Econometrics*, 201:144–169, 2017.
- [33] H. Ichimura and T. S. Thompson. Maximum likelihood estimation of a binary choice model with random coefficients of unknown distribution. *Journal of Econometrics*, 86:269–295, 1998.

- [34] G. W. Imbens and J. D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62:467–475, 1994.
- [35] C. Kleiber and J. Stoyanov. Multivariate distributions and the moment problem. *Journal of Multivariate Analysis*, 113:7–18, 2013.
- [36] I. Kotlarski. On characterizing the Gamma and the normal distribution. *Pacific Journal of Mathematics*, 20:69–76, 1967.
- [37] G. D. Lin. On the moment problems. *Statistics and Probability Letters*, 35:85–90, 1997.
- [38] G. D. Lin. Recent developments on the moment problem. 2017. Preprint [arXiv:1703.01027v1](https://arxiv.org/abs/1703.01027v1).
- [39] S. Mandelbrojt. *Séries adhérentes, régularisation des suites, applications*. Gauthiers-Villars, 1952.
- [40] M. A. Masten. Random coefficients on endogenous variables in simultaneous equations models. *The Review of Economic Studies*, 85(2):1193–1250, 2017.
- [41] A. Meister. Deconvolving compactly supported densities. *Mathematical Methods of Statistics*, 16:63–76, 2007.
- [42] M. Morimoto. *Analytic functionals on the sphere*. American Mathematical Society, 1962.
- [43] A. E. Nussbaum. Quasi-analytic vectors. *Archiv der Mathematik*, 6:179–191, 1965.
- [44] A. G. Pakes, W. Hung, and J. Wu. Criteria for the unique determination of probability distributions by moments. *Australian and New Zealand Journal of Statistics*, 43:101–111, 2001.
- [45] H. L. Pedersen. On Krein’s theorem for indeterminacy of the classical moment problem. *J. Approx. Theory*, 95:90–100, 1998.
- [46] L. C. Petersen. On the relation between the multidimensional moment problem and the one-dimensional moment problem. *Mathematica Scandinavica*, 51:361–366, 1982.
- [47] W. Rudin. *Real and complex analysis*. McGraw-Hill, 1973.
- [48] W. Rudin. *Function theory in the unit ball of \mathbb{C}^n* . Springer, 1980.
- [49] H. Salzmann and K. Zeller. Singularitäten unendlich oft differenzierbarer funktionen. *Mathematische Zeitschrift*, 62:354–367, 1955.
- [50] J. Schmets and M. Valdivia. On the extent of (non) quasi-analytic classes. *Archiv der Mathematik*, 56:593–600, 1991.
- [51] B. V. Shabat. *Introduction to complex analysis: functions of several variables, Part 2*. AMS, 1992.
- [52] J. Shohat and J. Tamarkin. *The problem of moments*. American Mathematical Society, 1970.
- [53] J. Stoyanov. *Counterexamples in probability*. John Wiley & Sons, 1997.
- [54] J. Stoyanov. Krein condition in probabilistic moment problems. *Bernoulli*, 6:939–949, 2000.
- [55] J. Stoyanov and G. D. Lin. Hardy’s condition in the moment problem for probability distributions. *Theory Probab. Appl.*, 57:811–820, 2012.
- [56] P. Tauvel. *Analyse complexe pour la licence 3*. Dunod, 2006.
- [57] K. Train. *Discrete choice methods with simulation*. Cambridge, 2009.
- [58] E. Vytlacil. Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70:331–341, 2002.
- [59] Y. Wu, J. Fan, H.-G. Müller, et al. Varying-coefficient functional linear regression. *Bernoulli*, 16(3):730–758, 2010.

[60] R. Young. *An introduction to nonharmonic Fourier series*. Academic press, 2001.

APPENDIX

Notations. For a differentiable function f of real variables, $f^{(\mathbf{m})}$ denotes $\prod_{j=1}^p \frac{\partial^{m_j}}{\partial x_j^{m_j}} f$ and $\text{supp}(f)$ its support. Define, for $\mu, \nu \in \mathfrak{M}_c^*(\Omega)$, where Ω is a closed set of \mathbb{R}^p , the equivalence $\mu \sim \nu$ if $s_\mu(\mathbf{m}) = s_\nu(\mathbf{m})$ for all $\mathbf{m} \in \mathbb{N}_0^p$. $\text{Aff}(\mathbb{R}^p)$ for the group of invertible affine transformations of \mathbb{R}^p . Denote by $S'(\mathbb{S}^p)$ the dual of $C^\infty(\mathbb{S}^p)$ [see 24].

A.1. Complements.

A.1.1. *Analytic and quasi-analytic classes of Fourier transforms of complex measures.* Given $b \in \mathbb{R}_+$, and $\{M_m\} \in ((0, \infty]^p)^{\mathbb{N}_0}$ where $\{M_{m,k}\}$ for all $k \in [p]$ are log-convex sequences with $M_{0,k} = 1$, let us introduce the class of functions

$$(A.1) \quad C^{\{M_m\}}(b) = \left\{ f \in C^\infty(\mathbb{R}^p) : \exists c : \forall \mathbf{m} \in \mathbb{N}_0^p, \left\| f^{(\mathbf{m})} \right\|_{L^\infty(\mathbb{R}^p)} \leq cb^{|\mathbf{m}|} \prod_{k=1}^p M_{k,m_k} \right\}.$$

It consists of complex valued functions and is a vector space. We denote by $C^{\{M_m\}} = \bigcup_{b \in \mathbb{R}_+} C^{\{M_m\}}(b)$. Like $\mathcal{A}(S)$, $C^{\{M_m\}}$ is an algebra with respect to multiplication [see 47]. From Theorem 19.9 in [47], $C^{\{M_m\}}$ consists of the functions f such that there exists $\rho > 0$ such that f can be extended uniquely to $\mathcal{H}^\infty(\{z \in \mathbb{C} : |\text{Im}(z)| < \rho\})$. Note that if $f \in C^{\{M_m\}}(b)$ and $\underline{x} \in \mathbb{R}^p$, then the function obtained by fixing all variables but the k^{th} to those of \underline{x} defines a function of $C^{\{M_{k,m_k}\}}(b)$. Thus, products of sets of uniqueness of the spaces of functions of a single variable are sets of uniqueness of $C^{\{M_m\}}(b)$.

By the Hölder inequality, if $\|\cdot\|$ is a norm and $\mu \in \mathfrak{M}^*(\mathbb{R}^p)$, $(\int_{\mathbb{R}^p} \|\mathbf{x}\|^m d\mu(\mathbf{x}))_{m \in \mathbb{N}_0}$ is log-convex and $M_0 = \mu(\mathbb{R}^p)$. Denote by $M(x)$ the trace function of $\{M_m\}$, where, for all $m \in \mathbb{N}$, $M_m^c = \exp(\sup_{x \geq 0}(mx - M(x)))$ and $M(x) = \sup_{m \in \mathbb{N}}(mx - \log(M_m))$ [see 39]. The sequences $\{1_m\}$ and $\{M_m^c\}$ are log-convex. If $S = \mathbb{R}$ and $\underline{\lim}_{m \rightarrow \infty} M_m^{1/m} = 0$ then $C^{\{M_m\}}(S) = C^{\{0_m\}}(S)$ and if $0 < \underline{\lim}_{m \rightarrow \infty} M_m^{1/m} < \infty$ then $C^{\{M_m\}}(S) = C^{\{1_m\}}(S)$, else, if the terms in the sequence are positive, $C^{\{M_m\}}(S) = C^{\{M_m^c\}}(S)$.

Let us also introduce certain subsets of $\mathcal{P}_c(\Omega)$ of the form

$$\mathcal{P}_c^{\{M_m\}}(\Omega) = \left\{ \mu \in \mathfrak{M}_c^*(\Omega) : \exists c, b : \forall m \in \mathbb{N}_0, k \in [p], s_{|\cdot|, |\mu|}(me_k) \leq cb^m M_{k,m} \right\}.$$

Proposition A.1 shows that $\mathcal{F}[\mathcal{P}_c] \subseteq C$ for the classes \mathcal{P}_c and C used in the examples (E.1), (E.3) and (E.4) related to the restriction $\mathcal{R}_{L,7}$.

Proposition A.1. *Let $\Omega \subseteq \mathbb{R}$ be a closed set, $\mu \in \mathfrak{M}_c^*(\Omega)$, $f = \mathcal{F}[\mu]$, and $\{M_m\} \in (0, \infty]^{\mathbb{N}_0}$ be log-convex sequence.*

- (1) *If there exists an increasing function m such that, for all $r \in \mathbb{R}_+$, $\int_{\mathbb{R}} e^{r|x|} d|\mu|(x) \leq 2m(r)$, then f belongs to*

$$\mathcal{A}_0(\mathbb{C}) = \left\{ f \in \mathcal{A}(\mathbb{C}) : \forall r \geq 0, \max_{z:|z|=r} |f(z)| \leq 2m(r) \right\};$$

- (2) *Let $\mu \in \mathcal{P}_c^{\{M_m\}}(\mathbb{R})$, where $\{M_m\}$ satisfies (12) then $f \in C^{\{M_m\}}$, which is a quasi-analytic class.*

(3) Let $\mu \in \mathcal{P}_c^{\{M_m\}}(\mathbb{R}_+)$, where $\{M_m\}$ satisfies (11) then $f \in C^{\{M_m\}}$ and

$$C^{\{M_m\}} \cap \{\mathcal{F}[\mu], \mu \in \mathfrak{M}_c^*(\mathbb{R}_+)\}$$

is a quasi-analytic class.

In all the cases of Proposition A.1, f is characterized by $(f^m(0))_{m \in \mathbb{N}_0}$, hence μ is determinate in $\mathfrak{M}_c^*(\mathbb{R})$ by injectivity of the Fourier transform.

Proof of Proposition A.1. Consider (1). Use that there exists a complex Borel function g with $|g| = 1$ such that $d\mu = gd|\mu|$ and rewrite all integrals as integrals with respect to $|\mu|$. Let $r \in \mathbb{R}_+$ and $S_r = \{z \in \mathbb{C} : |z| \leq r\}$. For all $x \in \text{supp}(\mu)$, $z \in S_r \mapsto e^{izx}$ is holomorphic and, for all $z \in S_r$, $x \in \mathbb{R} \mapsto e^{izx}$ is $|\mu|$ -integrable. For all compact $K \subset S_r$, for all $z \in K$, we have, for all $x \in \text{supp}(\mu)$, $|e^{izx}| \leq e^{r|x|}$ which is $|\mu|$ -integrable. The rest follows by the same argument as those of the proof of theorem p91 of [56] for complex variables.

Consider (2). Because $\mu \in \mathfrak{M}_c^*(\mathbb{R})$, we have $f \in C^\infty(\mathbb{R})$ but also clearly $f \in C^{\{M_m\}}$ because, for all $m \in \mathbb{N}_0$ and $x \in \mathbb{R}$, $|f^{(m)}(x)| \leq s_{|\cdot|, |\mu|}(m)$. The Denjoy-Carleman theorem [see, e.g., Theorem 19.11 in 47] yields that $C^{\{M_m\}}$ is quasi-analytic. Moreover, using that for all $m \in \mathbb{N}_0$, $f^{(m)}(0) = i^m s_\mu(m)$, μ is determinate in $\mathfrak{M}_c^*(\mathbb{R})$.

Consider (3), we adapt the proof of Theorem 4.1 in [12]. Define the measure μ_1 on \mathbb{R} by $d\mu_1(t) = d\mu(t^2)$. Thus, we have, for all $m = 2n$ with $n \in \mathbb{N}_0$, $s_{|\cdot|, |\mu_1|}(m) = 2s_{|\cdot|, |\mu|}(n)$ hence $\sum_{m \in \mathbb{N}} 1/(s_{|\cdot|, |\mu_1|}(2m))^{1/(2m)} = \sum_{m \in \mathbb{N}} 1/(2s_{|\cdot|, |\mu|}(m))^{1/(2m)}$ and $\sum_{m \in \mathbb{N}} 1/(s_{|\cdot|, |\mu|}(m))^{1/(2m)} \geq \sum_{m \in \mathbb{N}} 1/M_m^{1/(2m)} = \infty$. Hence, applying (2) to μ_1 , μ is determinate in $\mathfrak{M}_c^*(\mathbb{R}_+)$. Because we have, for all $(m, x) \in \mathbb{N}_0 \times \mathbb{R}$,

$$\left| \mathcal{F}^{(m)}[\mu](x) \right| = \frac{1}{2} \left| \mathcal{F}^{(2m)}[\mu_1](x) \right| \leq 2^{2m-1} s_{|\cdot|, |\mu_1|}(2m) \leq 2^{2m} s_{|\cdot|, |\mu|}(m),$$

$\mathcal{F}[\mu] \in C^{\{M_m\}}$, hence $\mathcal{F}^{(m)}[\mu](0) = i^m s_\mu(m)$ yields that

$$C^{\{M_m\}} \cap \{\mathcal{F}[\mu], \mu \in \mathfrak{M}_c^*(\mathbb{R}_+)\}$$

is quasi-analytic. □

A.1.2. *Criteria for determinacy.* When $\mu \in \mathfrak{M}^*(\mathbb{R}^{p+1})$, the determinacy of μ can be assessed by checking the Cramér condition:

$$(A.2) \quad \exists \rho > 0 : \forall \mathbf{t} : |\mathbf{t}|_2 \leq \rho, \int_{\mathbb{R}^{p+1}} e^{\mathbf{t}^\top \mathbf{x}} d\mu(\mathbf{x}) < \infty.$$

When $\mu \in \mathfrak{M}_1^*(\mathbb{R}^{p+1})$, this means that the moment generating function exists for small $|\mathbf{t}|_2$. Using that $\int_{\mathbb{R}^{p+1}} \cosh(|\mathbf{t}^\top \mathbf{x}|) d\mu(\mathbf{x}) = \int_{\mathbb{R}^{p+1}} \cosh(\mathbf{t}^\top \mathbf{x}) d\mu(\mathbf{x}) < \infty$, (A.2) is equivalent to:

$$(A.3) \quad \exists \rho > 0 : \forall \mathbf{t} : |\mathbf{t}|_2 \leq \rho, \int_{\mathbb{R}^{p+1}} e^{|\mathbf{t}^\top \mathbf{x}|} d\mu(\mathbf{x}) < \infty.$$

The Cramér condition rules out laws with heavy tails and holds for example for multivariate normals and skew-normals (see [35] and references therein). For a random variable X such that $\mathbb{S}_X \subseteq \mathbb{R}_+$, it is possible to use the weaker Hardy condition $\mathbb{E}[e^{t\sqrt{X}}] < \infty$ for small $|t|$ to assess determinacy (see [26, 27, 55]). We now give a simple extension of Theorem 2 in [38] to $\mathfrak{M}^*(\mathbb{R})$.

Proposition A.2. *Let $\mu \in \mathfrak{M}^*(\mathbb{R})$ with $\text{supp}(\mu) = \mathbb{R}_+$, then $\overline{\lim}_{m \rightarrow \infty} s_\mu(m+1)/(m^2 s_\mu(m)) < \infty$ implies the following equivalent statements:*

(P.A.2.a) $\exists t > 0 : \int_{\mathbb{R}_+} e^{t\sqrt{x}} d\mu(x) < \infty;$

(P.A.2.b) $\overline{\lim}_{m \rightarrow \infty} s_\mu(m)^{1/(2m)}/m < \infty;$

(P.A.2.c) $\overline{\lim}_{m \rightarrow \infty} (s_\mu(m))^{1/m}/(2m)! < \infty;$

which imply

$$(A.4) \quad \sum_{m \in \mathbb{N}} \frac{1}{s_\mu(m)^{\frac{1}{m}}} = \infty,$$

hence that μ is determinate.

(A.4) is the Carleman condition of the Stieltjes moment problem.

Corollary A.1 (Theorem 1 in [38]). *Let $\mu \in \mathfrak{M}^*(\mathbb{R})$, then $\overline{\lim}_{m \rightarrow \infty} s_\mu(2(m+1))/(m^2 s_\mu(2m)) < \infty$ implies the following equivalent statements:*

(C.A.1.a) $\exists t \in \mathbb{R} : \int_{\mathbb{R}_+} e^{tx} d\mu(x) < \infty;$

(C.A.1.b) $\exists t > 0 : \int_{\mathbb{R}_+} e^{t|x|} d\mu(x) < \infty;$

(C.A.1.c) $\overline{\lim}_{m \rightarrow \infty} s_\mu(2m)^{1/(2m)}/2m < \infty;$

(C.A.1.d) $\overline{\lim}_{m \rightarrow \infty} (s_\mu(2m)/(2m)!)^{1/m} < \infty;$

(C.A.1.e) $\overline{\lim}_{m \rightarrow \infty} (s_{|\cdot|, \mu}(m)/m)^{1/m} < \infty;$

which imply

$$(A.5) \quad \sum_{m \in \mathbb{N}} \frac{1}{s_\mu(2m)^{\frac{1}{m}}} = \infty,$$

hence that μ is determinate.

(A.5) is the Carleman condition of the Hamburger moment problem. Corollary A.1 (see also Corollary A.2 below) allows the $\overline{\lim}$ in (C.A.1.c) and (C.A.1.d) to be taken over the subsequence of even integers. (A.5) when the support is known to lie in a half line is weaker than (A.4) and nonbinding when $\text{supp}(\mu)$ is compact. Neither are necessary for determinacy (see [30] p93, [53] p113).

Integral criteria for $\mu \in \mathfrak{M}_1^*(\mathbb{R})$ with a density f can be easier to check than (A.4) and (A.5). Due to Theorem 3 in [46], they can be used on marginals to assess determinacy of multivariate measures. We have the following results from [37]:

- (i) When f is positive and the Krein condition $\int_{\mathbb{R}} -\log(f(x))/(1+x^2)dx < \infty$ holds, μ is not determinate; while, if f is also even, differentiable, and there exists $x_0 > 0$ such that, for $x \geq x_0 > 0$ and $x \mapsto f(x)$ decreases to 0, $x \mapsto -xf'(x)/f(x)$ increases to infinity (so-called Lin conditions), and $\int_{\mathbb{R}} -\log(f(x))/(1+x^2)dx = \infty$ then (A.4) holds.
- (ii) When $\text{supp}(\mu) \subseteq \mathbb{R}^+$, $f(x)$ is positive on $[x_0, \infty)$ for $x_0 > 0$, and the Krein condition $\int_{x_0}^{\infty} -\log(f(x^2))/(1+x^2)dx < \infty$ holds, μ is not determinate (see [44] and [45] for weaker conditions on the interval where f is positive); while, if f is positive and differentiable on \mathbb{R}_+ , and, for $x \geq x_0$, $x \mapsto f(x)$ decreases to 0 and $x \mapsto -xf'(x)/f(x)$ increases to infinity (so-called Lin conditions), and $\int_{\mathbb{R}_+} -\log(f(x^2))/(1+x^2)dx = \infty$, then (A.5) holds.

The Krein condition is not necessary for $\mu \in \mathfrak{M}_1^*(\mathbb{R})$ to be indeterminate (see [53] p114) in the Hamburger and Stieltjes cases.

Theorem 3 in [46] and Proposition A.1 imply the following results for multivariate measures.

Corollary A.2. *Let $q \in \{0, 1, \dots, p+1\}$, $\mu_1, \mu_2 \in \mathfrak{M}_c^*(\mathbb{R}^{p+1})$ such that for known $\mathfrak{B} \in GL(\mathbb{R}^{p+1})$ $\text{supp}(\mathfrak{B}_* \mu_1) = \text{supp}(\mathfrak{B}_* \mu_2) = \mathbb{R}^q \times \mathbb{R}_+^{p+1-q}$ and $\mu_1 \sim \mu_2$. We have $\mu_1 = \mu_2$ if for $\mathfrak{A} \in \text{Aff}(\mathbb{R}^{p+1})$ leaving invariant $\text{span}\{e(q+1), \dots, e(p+1)\}$*

$$\begin{aligned} \min_{j=1, \dots, q} \sum_{m \in \mathbb{N}_0} \frac{1}{(s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_1|(2me(j)) + s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_2|(2me(j)))^{\frac{1}{2m}}} &= \infty; \\ \min_{j=q+1, \dots, p+1} \sum_{m \in \mathbb{N}_0} \frac{1}{(s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_1|(me(j)) + s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_2|(me(j)))^{\frac{1}{2m}}} &= \infty; \end{aligned}$$

or, when $\mu_1, \mu_2 \in \mathfrak{M}^*(\mathbb{R}^{p+1})$,

$$(A.6) \quad \min_{j=1, \dots, q} \sum_{m \in \mathbb{N}_0} \frac{1}{s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_1|(2me(j))^{\frac{1}{2m}}} = \infty;$$

$$(A.7) \quad \min_{j=q+1, \dots, p+1} \sum_{m \in \mathbb{N}_0} \frac{1}{s_{|\mathfrak{A} \circ \mathfrak{B}} * \mu_1|(me(j))^{\frac{1}{2m}}} = \infty;$$

where, in both cases, if $q = 0$, we remove the first minimum, while, if $q = p$, we remove the second minimum.

Proof of Corollary A.2. When $q = p+1$, the statement is obtained by applying the Denjoy-Carleman theorem to $\mathcal{F}[(\mathfrak{A} \circ \mathfrak{B})_*(\mu_1 - \mu_2)]$ (see Theorem 3.2 in [12] for the case of complex measures and else Theorem 2.3 in [16]). When $q < p$, we apply this result to $(\mathfrak{C} \circ \mathfrak{B})_* \mu_j$ for $j = 1, 2$, where $\mathfrak{C}((\mathbf{x}_1, \dots, \mathbf{x}_{p+1})^\top) = (\mathbf{x}_1, \dots, \mathbf{x}_q, \sqrt{\mathbf{x}_{q+1}}, \dots, \sqrt{\mathbf{x}_{p+1}})$ and proceed like in the proof of Theorem 4.1 in [12]. \square

The Denjoy-Carleman theorem applied to $\mathcal{F}[(\mathfrak{A} \circ \mathfrak{B})_* \mu]$ where $\mu \in \mathfrak{M}_c(\mathbb{R}^{p+1})$ only requires assumptions on marginals, hence the conditions in Corollary A.2 only involve these marginals, this is in line with Theorem 3 in [46]. [43] (see also Theorem 5 in [4]) states this result for elements of $\mathfrak{M}^*(\mathbb{R}^{p+1})$ when $q = p+1$ without allowing for affine transformations. Due to Proposition 3.10 in [17], it is possible that the sum of two functions belonging to classes defined by two different $p+1$ -tuples of sequences which both satisfy the sufficient condition of Denjoy-Carleman theorem do not belong to such a class. Rather than (A.6)-(A.7), also called Carleman conditions, [52] assumes the stronger condition

$$(A.8) \quad \sum_{m \in \mathbb{N}_0} \frac{1}{\left(\int_{\mathbb{R}^{p+1}} \sum_{j=1}^{p+1} \mathbf{x}_j^{2m} d\mu(\mathbf{x}) \right)^{\frac{1}{2m}}} = \infty.$$

Corollary A.3. *Let $q \in \{0, 1, \dots, p+1\}$, $\mu_1, \mu_2 \in \mathfrak{M}_c^*(\mathbb{R}^{p+1})$ such that for known $\mathfrak{B} \in GL(\mathbb{R}^{p+1})$ $\text{supp}(\mathfrak{B}_* \mu_1) = \text{supp}(\mathfrak{B}_* \mu_2) = \mathbb{R}^q \times \mathbb{R}_+^{p+1-q}$ and $\mu_1 \sim \mu_2$.*

$\mu_1 = \mu_2$ if there exist functions W_j for $j = 1, \dots, p+1$ from \mathbb{R} to $[0, \infty]$ such that:

$$(A.9) \quad \min \left\{ \min_{j=1, \dots, q} \sum_{m \in \mathbb{N}} \frac{1}{\left\| x \mapsto \frac{x^{2m}}{W_j(\sqrt{x})} \right\|_{L^\infty(\mathbb{R})}^{\frac{1}{2m}}}, \min_{j=q+1, \dots, p+1} \sum_{m \in \mathbb{N}} \frac{1}{\left\| x \mapsto \frac{x^{2m}}{W_j(\sqrt{x})} \right\|_{L^\infty(\mathbb{R}_+)}^{\frac{1}{2m}}} \right\} = \infty;$$

where, if $q = 0$, we remove the first minimum, while, if $q = p$, we remove the second minimum; and $\mathfrak{A} \in \text{Aff}(\mathbb{R}^{p+1})$ leaving invariant $\text{span}\{\mathbf{e}(q+1), \dots, \mathbf{e}(p+1)\}$ such that:

$$\begin{aligned} & \sum_{j=1, \dots, p+1} \int_{\mathbb{R}} W_j(\mathbf{x}_j) d(\Pi[\mathbf{e}(j)]_* (|\mathfrak{A} \circ \mathfrak{B}|_* \mu_1 + |\mathfrak{A} \circ \mathfrak{B}|_* \mu_2))(\mathbf{x}) < \infty; \\ \text{or} \quad & \sum_{j=1, \dots, p+1} \int_{\mathbb{R}} W_j(\mathbf{x}_j) d((\Pi[\mathbf{e}(j)] \circ \mathfrak{A} \circ \mathfrak{B})_* \mu_1)(\mathbf{x}) < \infty \text{ if } \mu_1, \mu_2 \in \mathfrak{M}^*(\mathbb{R}^{p+1}). \end{aligned}$$

Moreover, $\mu_1 = \mu_2$ if there exist functions W_j for $j = 1, \dots, p+1$ from \mathbb{R} to $[c_j, \infty]$, where $c_j > 0$, satisfying (A.9), $\mathfrak{A} \in \text{Aff}(\mathbb{R}^{p+1})$ leaving invariant $\text{span}\{\mathbf{e}(q+1), \dots, \mathbf{e}(p+1)\}$, and a function W from \mathbb{R}^{p+1} to $[0, \infty]$ such that, for $|\mathbf{x}|_2$ large enough, $W(\mathfrak{A} \circ \mathfrak{B}(\mathbf{x})) \geq \prod_{j=1}^{p+1} W_j(\mathbf{x}_j)$, with $\int_{\mathbb{R}^{p+1}} W(\mathfrak{A} \circ \mathfrak{B}(\mathbf{x})) d(|\mu_1| + |\mu_2|)(\mathbf{x}) < \infty$ or $\int_{\mathbb{R}^{p+1}} W(\mathfrak{A} \circ \mathfrak{B}(\mathbf{x})) d\mu_1(\mathbf{x}) < \infty$ if $\mu_1, \mu_2 \in \mathfrak{M}^*(\mathbb{R}^{p+1})$.

Proof of Corollary A.3. Let us prove the moreover part of the statement. The first part is obtained in a similar manner. Assuming that W_j are bounded from below by $c_j > 0$ for $j = 1, \dots, p+1$, we use that for $j \in \{1, \dots, p+1\}$, $m \in \mathbb{N}$, and $\mu, \nu \in \mathfrak{M}_c^*(\mathbb{R}^{p+1})$, we have:

$$\begin{aligned} s_{|\mu|}(m\mathbf{e}(j)) + s_{|\nu|}(m\mathbf{e}(j)) & \leq \|\mathbf{x}^m / W_j(\mathbf{x})\|_{L^\infty(\mathbb{R})} \int_{\mathbb{R}^{p+1}} W_j(\mathbf{x}_j) d(|\mu| + |\nu|)(\mathbf{x}) \\ & \leq \frac{1}{\prod_{\substack{k=1, \dots, p+1 \\ k \neq j}} c_k} \|\mathbf{x}^m W_j(\mathbf{x})^{-1}\|_{L^\infty(\mathbb{R})} \int_{\mathbb{R}^{p+1}} W(\mathbf{x}) d(|\mu| + |\nu|)(\mathbf{x}). \end{aligned}$$

We use the same argument for $\mu_1 \in \mathfrak{M}^*(\mathbb{R}^{p+1})$. We conclude using Corollary A.2. \square

The first sufficient condition in Corollary A.3 is the less demanding. Due to Theorem 3.14 in [17], if W is a function from \mathbb{R} to $[c, \infty]$ for $c > 0$, the next statements are equivalent:

- (QAW.1) $\sum_{m \in \mathbb{N}} 1 / \left\| x \mapsto \frac{x^{2m}}{W(x)} \right\|_{L^\infty(\mathbb{R})}^{1/(2m)} = \infty$;
- (QAW.2) There exists a function \tilde{W} with values in $[0, \infty)$ and $R > 0$ such that, for $|x| > R$, $W(x) \geq \tilde{W}(x)$ and $\tilde{W}(x) = \tilde{W}(-x) < \infty$, $x \mapsto \log(\tilde{W}(e^x))$ is convex on $(\log(R), \infty)$ and $\int_R^\infty \log \tilde{W}(s) / (1 + s^2) ds = \infty$;
- (QAW.3) There exist $\epsilon > 0$, and a nondecreasing nonnegative function $\rho \in C^\infty([0, \infty))$ equal to 0 on $[0, \epsilon]$, such that, for all $x \in \mathbb{R}$, $W(x) \geq \exp\left(\int_0^{|x|} (\rho(s)/s) ds\right)$ and $\int_0^\infty \rho(s) / (1 + s^2) ds = \infty$.

For example, if there exist $R \in \mathbb{R}$, $(a_j)_{j \in \mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0}$, and $(p_j)_{j \in \mathbb{N}_0} \in (-\infty, 1]^{\mathbb{N}_0}$ equal to 0 for j large enough, such that

$$W(x) \geq \exp\left(\frac{x^2}{\prod_{j=0}^{\infty} \log_j^{p_j}(a_j|x|)}\right) \mathbb{1}\{|x| \geq R\} \quad \left(\text{e.g., } W(x) = \exp\left(\frac{|x|}{a_0 \log(a_1|x|)}\right) \mathbb{1}\{|x| \geq R\}\right)$$

then W satisfies (QAW.3). Replacing x by \sqrt{x} and using that the above lower bound is even, we obtain that $\sum_{m \in \mathbb{N}} 1/\|x \mapsto x^{2m}/W(\sqrt{x})\|_{L^\infty(\mathbb{R}_+)}^{1/(2m)} = \infty$ if there exist $(C, R) \in \mathbb{R}^2$, $(a_j)_{j \in \mathbb{N}_0} \in (0, \infty)^{\mathbb{N}_0}$, and $(p_j)_{j \in \mathbb{N}_0} \in (-\infty, 1]^{\mathbb{N}_0}$ equal to 0 for j large enough, such that

$$W(x) \geq C \exp\left(\frac{x}{\prod_{j=0}^{\infty} \log_j^{p_j}(a_j\sqrt{x})}\right) \mathbb{1}\{x \geq R\} \quad \left(\text{e.g., } W(x) = \exp\left(\frac{\sqrt{x}}{a_0 \log(a_1x)}\right) \mathbb{1}\{x \geq R\}\right).$$

Because $\sum_{m \in \mathbb{N}} 1/\|x \mapsto x^{2m}/W(x)\|_{L^\infty(\mathbb{R})}^{1/(2m)} = \infty$ implies $\sum_{m \in \mathbb{N}} 1/\|\mathbf{x} \mapsto (\mathbf{x}^\top \mathbf{f}(j))^{2m}/W(|\mathbf{x}|_2)\|_{L^\infty(\mathbb{R}^q)}^{1/(2m)} = \infty$ for all bases $(\mathbf{f}(j))_{j=1}^q$ of \mathbb{R}^q and similarly for \mathbb{R}_+^{p+1-q} , assuming that \mathfrak{B} is identity, we can replace (A.9) by the stronger condition (not involving \mathfrak{A})

$$W(\mathbf{x}) \geq \widetilde{W}\left(|(\mathbf{x}_1, \dots, \mathbf{x}_q)^\top|_2\right) \widetilde{W}\left(\sqrt{|(\mathbf{x}_{q+1}, \dots, \mathbf{x}_{p+1})^\top|_2}\right),$$

where \widetilde{W} is a function from \mathbb{R} to $[c, \infty]$ for $c > 0$ which satisfies the equivalent (QAW.1)-(QAW.3). The function $1/W$ is called a quasi-analytic weight in [16, 17].

The approach using quasi-analytic weights is closely related to the Krein and Lin conditions. Assume that \mathfrak{A} and \mathfrak{B} are the identity, $q = p + 1$ (the remaining cases can be treated similarly), and $\mu \in \mathfrak{M}^*(\mathbb{R}^{p+1})$ has marginals $\Pi[e(j)]_* \mu$ with densities $f_j(\mathbf{x}_j)$ for $j = 1, \dots, p+1$ which are even and positive and such that $-\mathbf{x}_j f'_j(\mathbf{x}_j)/f_j(\mathbf{x}_j) - \alpha_j \log(\log(\mathbf{x}_j))$ with $\alpha_j > 1$ is increasing for $|\mathbf{x}_j| \geq x_{0j} \geq 0$. Then, for $j = 1, \dots, p+1$, $W_j(\mathbf{x}_j) = 1/\max(f_j(\mathbf{x}_j) \log(|\mathbf{x}_j|^{\alpha_j}, 1)$ satisfies (QAW.2) and $\int_{\mathbb{R}} W_j(\mathbf{x}_j) f_j(\mathbf{x}_j) d\mathbf{x}_j < \infty$, hence the conclusions of Corollary A.3 hold. The advantage of the approach using quasi-analytic weights is that, unlike the Krein and Lin conditions, it is not required that a density exists.

A.1.3. Complements on the sets of uniqueness. Examples of sets of uniqueness for analytic classes are given in the examples (E.1) and (E.2). We only give here more details about (E.1).

Let $m \in \mathbb{R}_+^{[0, \infty)}$, $\lim_{r \rightarrow \infty} m(r) = \infty$, and

$$\mathcal{A}_0(\mathbb{C}) = \left\{ f \in \mathcal{A}(\mathbb{C}) : \exists C > 0 : \forall r \geq 0, \max_{z:|z|=r} |f(z)| \leq C m(r) \right\}.$$

U is a set of uniqueness of such a class $\mathcal{A}_0(\mathbb{C})$ if, for $\epsilon > 0$, we have

$$(A.10) \quad \exists \alpha > 1 : \overline{\lim}_{r \rightarrow \infty} \frac{\log(\alpha)}{\log(m(\alpha r))} |U \cap ((-r, r) \setminus \{0\})| > 1.$$

Hence, tU is a set of uniqueness for all $t \neq 0$ if

$$\forall t > 0, \exists \alpha > 1 : \overline{\lim}_{r \rightarrow \infty} \frac{\log(\alpha)}{\log(m(\alpha t r))} |U \cap ((-r, r) \setminus \{0\})| > 1.$$

A sufficient condition when $m(r) = e^{\rho r}$, which occurs when f is the Fourier transform of the difference of two complex measures supported in $[-\rho, \rho]$, is

$$\overline{\lim}_{r \rightarrow \infty} \frac{1}{r} |U \cap ((-r, r) \setminus \{0\})| = \infty.$$

Let us explain the above characterization. If $f \in \mathcal{A}_0(\mathbb{C})$, there exists $k \in \mathbb{N}$ such that $f(z) = z^k g(z)$ and $g(0) \neq 0$. Jensen's formula applied to g (see 15.18 in [47]), with $\max_{z:|z|=r} |g(z)| = \max_{z:|z|=r} |f(z)|/z^k \leq Cm(r)/r^k$, ensures

$$|g(0)| \prod_{k=1}^{n_g(\alpha r)} \frac{\alpha r}{|\omega_k|} \leq \frac{Cm(\alpha r)}{\alpha^k r^k},$$

where $\omega_1, \dots, \omega_{n_g(\alpha r)}$ are the zeros of g in $B(0, \alpha r)$ ranked according to their multiplicity and $n_g(\alpha r)$ is the number of zeros of g with multiplicity in $B(0, \alpha r)$. Thus, we have

$$\frac{Cm(\alpha r)}{\alpha^k r^k} \geq |g(0)| \prod_{k=1}^{n_g(\alpha r)} \frac{\alpha r}{|\omega_k|} \geq |g(0)| \alpha^{n_g(\alpha r)},$$

hence $n_g(r) \leq \log(Cm(\alpha r)(\alpha r)^{-k} / |g(0)|) / \log(\alpha)$. We have

$$\frac{n_g(r) \log(\alpha)}{\log(m(\alpha r))} \leq 1 + \frac{\log(C(\alpha r)^{-k} / |g(0)|)}{\log(m(\alpha r))}$$

then $\overline{\lim}_{r \rightarrow \infty} n_g(r) \log(\alpha) / \log(m(\alpha r)) \leq 1$. This yields that U satisfying (A.10) is a set of uniqueness of $\mathcal{A}_0(\mathbb{C})$.

The following sets U are sets of uniqueness of the following quasi-analytic and V -quasi-analytic classes and complement (E.4).

Example 1. V is a subset of the interior of U , which is nonempty. Indeed, a function which is zero on U has all its partial derivatives or Laplacians (for functions on $S \subseteq \mathbb{S}^p$) equal to zero at every point in V .

Example 2. [Theorem 4b in 31] Let $M_m = \nu(m)^m m!$, where $x \in [0, \infty) \mapsto \nu(x)$ is increasing and continuously differentiable such that $\nu(0) = 1$ and

$$\lim_{x \rightarrow \infty} \frac{x\nu'(x)}{\nu(x)} = 0.$$

U is a set of uniqueness of $C^{\{M_m\}}(b)$ (which contains non-analytic functions) if

$$(A.11) \quad \overline{\lim}_{r \rightarrow \infty} \frac{1}{r} \int_1^{|U \cap (-r, r)|} M(\log(u)) \frac{du}{u^2} > \frac{\pi b}{2},$$

where $M(x)$ is the *trace function* of $\{M_m\}$ defined by $M(x) = \sup_{m \in \mathbb{N}} (mx - \log(M_m))$ (see [39]).

A.1.4. $\mathcal{A}(S)$ is a small subset of $C^\infty(S)$. Proposition A.3 can be found in [50], it is of the same spirit as the statement that the complement of nowhere analytic functions in $C^\infty(S)$ is meager with empty interior, which we detail below. $C^{\{M_m\}}(S)$ is defined like $C^{\{M_m\}}$ but we replacing \mathbb{R}^p by an open set S .

Proposition A.3. *If the condition (12) fails, then $C^{\{M_m\}}(S)$ contains a vector space of dimension 2^{\aleph_0} such that for all nonzero element f there does not exist $x \in S$, a neighborhood \mathcal{N}_x of x , and a sequence $\{M_{x,m}\}$ such that f belongs to a quasi-analytic class $C^{\{M_{x,m}\}}(\mathcal{N}_x)$.*

Assume, for simplicity, that $S \subset \mathbb{R}^p$ is compact and as usual $C^\infty(S)$ is equipped with the distance $d(f, g) = \sum_{\mathbf{m} \in \mathbb{N}_0^p} \min\left(1/2^{\mathbf{m}}, \|f^{(\mathbf{m})} - g^{(\mathbf{m})}\|_{L^\infty(S)}\right)$. By the Cauchy-Hadamard theorem [see 51], $f \in C^\infty(S)$ is analytic at $\underline{\mathbf{x}}$ if and only if there exist $b, c \in \mathbb{N}$ such that f belongs to

$$T(\underline{\mathbf{x}}, b, c) = \left\{ f \in C^\infty(S) : \forall \mathbf{m} \in \mathbb{N}_0^p, \left| f^{(\mathbf{m})}(\underline{\mathbf{x}}) \right| \leq cb^{|\mathbf{m}|} \mathbf{m}! \right\},$$

thus $\mathcal{A}(S) = \bigcup_{\underline{\mathbf{x}} \in \mathbb{Q}^p \cap S, b, c \in \mathbb{N}} T(\underline{\mathbf{x}}, b, c)$. This is a countable union of closed sets, hence closed. Moreover, $T(\underline{\mathbf{x}}, b, c)$ has empty interior. Indeed [see 49], given f in $T(\underline{\mathbf{x}}, b, c)$, for all $\epsilon > 0$, $m \in \mathbb{N}$ such that $\sum_{j=m}^\infty 2^{-j} < \epsilon$, $c < (\epsilon b^m / (2m!))^{1/(2m)}$, the functions $f_\epsilon : \mathbf{x} \in S \mapsto f(\mathbf{x}) + \epsilon \cos(c(\mathbf{x}_1 - \underline{\mathbf{x}}_1)) / (2c^m)$ are such that $d(f_\epsilon, f) < \epsilon$ and $\left| f_\epsilon^{(2m)}(\underline{\mathbf{x}}) - f^{(2m)}(\underline{\mathbf{x}}) \right| > b^{2m} (2m)!$, hence $f_\epsilon \notin T(\underline{\mathbf{x}}, b, c)$. Due to Baire's theorem, the meager set $\mathcal{A}(S)$ of the complete metric space $C^\infty(S)$ has an empty interior. With the arguments in [11], the complement of nowhere analytic functions in $C^\infty(S)$, hence containing $\mathcal{A}(S)$, can be shown to be meager with empty interior.

A.2. Proofs.

Proof of Theorem 1 under $\mathcal{R}_{L,1}$. Assume that \mathbb{P}_Γ and \mathbb{P}_Γ^* both give rise to the same collection $\left(\mathbb{P}_{(1, \mathbf{x}^\top) \Gamma} \right)_{\mathbf{x} \in \mathbb{S}_\mathbf{X}}$. By (ii), we have for all $\mathbf{x} \in \mathbb{S}_\mathbf{X}$ and t in \mathbb{R} ,

$$(A.12) \quad \varphi_\alpha(t) \prod_{j=1}^p \varphi_{\beta_j}(t\mathbf{x}_j) = \varphi_\alpha^*(t) \prod_{j=1}^p \varphi_{\beta_j}^*(t\mathbf{x}_j).$$

Because the value at 0 of a characteristic function is 1, taking $\mathbf{x} = \mathbf{0}_c$, yields $\varphi_\alpha = \varphi_\alpha^*$. Because φ_α^* is continuous at 0, for all $\mathbf{x} \in \mathbb{S}_\mathbf{X}$ and $t \in (-t_0, t_0)$ for t_0 small enough,

$$(A.13) \quad \prod_{j=1}^p \varphi_{\beta_j}(t\mathbf{x}_j) = \prod_{j=1}^p \varphi_{\beta_j}^*(t\mathbf{x}_j).$$

Taking $\mathbf{x} = \mathbf{x}(1)$ yields, for all $t \in (-t_0, t_0)$, $\varphi_{\beta_1}(\mathbf{x}(1)_1 t) = \varphi_{\beta_1}^*(\mathbf{x}(1)_1 t)$, hence $\varphi_{\beta_1}(t) = \varphi_{\beta_1}^*(t)$ for all $t \in (-t_0/\mathbf{x}(1)_1, t_0/\mathbf{x}(1)_1)$. Hence \mathbb{P}_{β_1} and $\mathbb{P}_{\beta_1}^*$ have same moments, so $\mathbb{P}_{\beta_1} = \mathbb{P}_{\beta_1}^*$. We conclude by iterating this procedure. \square

Proof of Theorem 1 under $\mathcal{R}_{L,2}$. By (A.12), (A.13) holds for all $\mathbf{x} \in \mathbb{S}_\mathbf{X}$ and t in the complement of the zeros of φ_α^* . Because the complement of the set of zeros of φ_α^* is dense and $t \mapsto \prod_{j=1}^p \varphi_{\beta_j}(t\mathbf{x}_j)$ and $t \mapsto \prod_{j=1}^p \varphi_{\beta_j}^*(t\mathbf{x}_j)$ are continuous, (A.13) holds for all $\mathbf{x} \in \mathbb{S}_\mathbf{X}$ and $t \in \mathbb{R}$. Taking $\mathbf{x} = \mathbf{x}(1)$ yields, for all $t \in \mathbb{R}$, $\varphi_{\beta_1}(\mathbf{x}(1)_1 t) = \varphi_{\beta_1}^*(\mathbf{x}(1)_1 t)$, so $\varphi_{\beta_1} = \varphi_{\beta_1}^*$. So, for all $\mathbf{x} \in \mathbb{S}_\mathbf{X}$ and $t \in \mathbb{R}$, $\prod_{j=2}^p \varphi_{\beta_j}(t\mathbf{x}_j) = \prod_{j=2}^p \varphi_{\beta_j}^*(t\mathbf{x}_j)$, and we conclude by iteration. \square

Proof of Theorem 1 under $\mathcal{R}_{L,3}$. Take $S = \mathbb{R}^{p+1}$ and $\mathfrak{F}(S)$ the vector space spanned by $\mathcal{F}[\mathcal{P}]$. An element of $\mathfrak{F}(S)$ is of the form $\mathcal{F}[\mu - \nu]$, where $\mu, \nu \in \mathcal{P}$. For all $k \in \mathbb{N}_0$ and

$\mathbf{x} \in \mathbb{S}_{\mathbf{X}}$,

$$(-i)^k \partial_t^{(k)} \mathcal{F}[\mu - \nu](t, t\mathbf{x}) \Big|_{t=0} = \int_{\mathbb{R}^{p+1}} \left(\gamma_1 + \sum_{j=1}^p \gamma_{j+1} \mathbf{x}_j \right)^k d(\mu - \nu)(\gamma) = P_k(\mathbf{x}),$$

where P_k are the polynomials with real coefficients

$$P_k(Z_1, \dots, Z_p) = \sum_{j \in \mathbb{N}_0^{p+1}: |j|_1 = k} \binom{k}{j_1, \dots, j_{p+1}} \left(\int_{\mathbb{R}^{p+1}} \gamma^j d(\mu - \nu)(\gamma) \right) Z_1^{j_2} \dots Z_p^{j_{p+1}}.$$

So, by (i), $\mathcal{F}[\mu - \nu] = 0$ on $U = \{t(1, \mathbf{x}) : t \in \mathbb{R}, \mathbf{x} \in \mathbb{S}_{\mathbf{X}}\}$ implies, for all $\mathbf{m} \in \mathbb{N}_0^{p+1}$, $s_\mu(\mathbf{m}) = s_\nu(\mathbf{m})$, hence $\mu = \nu$.

To prove the second statement, assume there exists $P \in \mathbb{R}[Z_1, \dots, Z_p]$ nonzero of degree $k \in \mathbb{N}$, such that, $\forall \mathbf{x} \in \mathbb{S}_{\mathbf{X}}$, $P(\mathbf{x}) = 0$. Let Q be the corresponding homogeneous polynomial in $\mathbb{R}[Z_1, \dots, Z_{p+1}]$ of degree k such that, $\forall \mathbf{x} \in \mathbb{S}_{\mathbf{X}}$, $Q(1, \mathbf{x}) = P(\mathbf{x})$. Let \mathbb{P}_{Γ}^* of support in $[-1, 1]^{p+1}$ and density with respect to the Lebesgue measure bounded from below by $c > 0$, and the infinitely differentiable function $q(\gamma) = \prod_{j=1}^{p+1} \gamma_j \exp(-1/(\gamma_j^2 - 1)) \mathbb{1}\{\gamma_j \in [-1, 1]\}$. Then, $\mathbb{P}_{\Gamma} = \mathbb{P}_{\Gamma}^* + ch/\|h\|_{\infty} d\gamma$, where $h = Q(\partial_1, \dots, \partial_{p+1})q$ is a probability. By properties of the Fourier transform of measures in $\mathfrak{M}^*(\mathbb{R}^{p+1})$ and homogeneity, we conclude from

$$\mathcal{F}[\mathbb{P}_{\Gamma}](t, t\mathbf{x}) = \mathcal{F}[\mathbb{P}_{\Gamma}^*](t, t\mathbf{x}) + t^k Q(1, \mathbf{x}) \mathcal{F}[h](t, t\mathbf{x}).$$

□

Proof of Proposition 1. Start by proving (P6.1). Take $J \subseteq \mathbb{Z} \setminus \{0\}$ finite, $J_s \subseteq J$ the positive indices j in J such that $-j \in J$.

Let $(b_j)_{j \in J} \in \mathbb{N}_0^{|J|}$ such that $\sum_{j \in J} b_j \lambda_j = 0$. We have

$$\sum_{j \in J} b_j \lambda_j = \sum_{j \in J_s} \frac{b_j + b_{-j}}{4r^j} + \sum_{j \in J \setminus (J_s \cup -J_s)} \frac{b_j}{4r^{|j|}} + \sum_{j \in J} j b_j,$$

hence, $P(r) = 0$, where $j_0 = \max_{j \in J} |j|$ and

$$P(X) = \sum_{j \in J_s} (b_j + b_{-j}) X^{j_0 - j} + \sum_{j \in J \setminus J_s} b_j X^{j_0 - |j|} + 4X^{j_0} \sum_{j \in J} j b_j \in \mathbb{Z}[X].$$

Because r is transcendental over \mathbb{Z} and $0 \notin J$, for all $j \in J \setminus (J_s \cup -J_s)$, $b_j = 0$, and, for all $j \in J_s$, $b_j + b_{-j} = 0$ thus $b_j = b_{-j} = 0$ because $b_j, b_{-j} \in \mathbb{N}_0$. Hence, for all $j \in J$, $b_j = 0$.

Now, take $P \in \mathbb{C}(X)[X_1, \dots, X_p]$, where $p = |J|$, which is zero when evaluated at $(f_j)_{j \in J}$. Hence, we have $\sum_k c_k(z) \exp(z \sum_{j \in J} b_{k,j} \lambda_j) = 0$, where the sum over k is finite, all $b_{k,j}$ belong to \mathbb{N}_0 , and c_k are rational functions. All exponentials in this sum are distinct by the above computations and we conclude that all c_k s are zero by taking limits in \mathbb{C} . This yields the result.

(P6.2) follows from Kadec's 1/4-Theorem [see Theorem 14 page 42 in 60] because $\sup_{j \in \mathbb{Z}} |\lambda_j - j| < 1/4$. □

Proof of Theorem 1 under $\mathcal{R}_{L,j}$ for $j \in \{4, 5\}$. Denote by $\langle, \rangle_{\ell_2(\mathbb{N})}$ the inner product in $\ell_2(\mathbb{N})$. Start by considering $\mathcal{R}_{L,4}$. By a similar argument as the one involving (3), we can consider without loss of generality $\{\underline{x}_m\} = \{0_m\}$ in $\mathcal{R}_{L,4}$ (i). Take $S = \ell_2(\mathbb{N})$ and $\mathfrak{F}(S)$ the vector space spanned by $\mathcal{F}[\mathcal{P}]$. An element of $\mathfrak{F}(S)$ is of the form $\mathcal{F}[\mu - \nu]$, where

$\mu, \nu \in \mathcal{P}$. Let $n \in \mathbb{N}$. We have, for all $\{z_m\} \in G_{n+1}$ and $t \in \mathbb{R}$, $\int_{\ell_2(\mathbb{N})} e^{it\langle \{\gamma_m\}, \{z_m\} \rangle_{\ell_2(\mathbb{N})}} d(\mu - \nu)(\{\gamma_m\}) = \int_{G_{n+1}} e^{it\langle \{\gamma_m\}, \{z_m\} \rangle_{\ell_2(\mathbb{N})}} d\Pi_{G_{n+1}*}(\mu - \nu)(\{\gamma_m\})$. We can now proceed like in the proof of Theorem 1 under $\mathcal{R}_{L,3}$ and obtain $\Pi_{G_{n+1}*}\mu = \Pi_{G_{n+1}*}\nu$ hence $\mathcal{F}[\Pi_{G_{n+1}*}\mu - \Pi_{G_{n+1}*}\nu] = 0$. Using that $\cup_{n \in \mathbb{N}} G_{n+1}$ is dense in $\ell_2(\mathbb{N})$ and the continuity of the Fourier transform we conclude $\mu = \nu$.

Identification under $\mathcal{R}_{L,5}$ is corollary of Theorem 4.1 in [15]. \square

Proof of Theorem 1 under $\mathcal{R}_{L,7}$. The proof relies on a weaker form of diagram (4).

Take $\mathbb{P}_{\Gamma} \in \mathcal{P}$ and $U_l = \prod_{k=1}^l V_k$ for all $l \in [p]$. For $\beta \in \mathbb{C}^p$ and $j \in [p]$, denote by $\beta_{[j]} = (\beta_1, \dots, \beta_j)^\top$. The function $\mathcal{F}[\mathbb{P}_{\Gamma} - \mathbb{P}_{\Gamma}^*]$ is 0 on U , so for $t \neq 0$ and $\mathbf{u} \in U_{p-1}$, $g_{p,t,\mathbf{u}} : z \in \mathbb{R} \mapsto \mathcal{F}[\mathbb{P}_{\Gamma} - \mathbb{P}_{\Gamma}^*](t, t\mathbf{u}, z) = \mathcal{F}[\mu_{p,t,\mathbf{u}}(\cdot)](z)$ is 0 on tV_p , where

$$\mu_{p,t,\mathbf{u}}(\cdot) = \int_{\mathbb{R}^p} e^{it(a + \mathbf{b}_{[p-1]}^\top \mathbf{u})} d(\mathbb{P}_{\Gamma} - \mathbb{P}_{\Gamma}^*)(a, \mathbf{b}_{[p-1]}, \cdot).$$

For h nonnegative and measurable, we have $\int_{\mathbb{R}} h(z) d|\mu_{p,t,\mathbf{u}}|(z) \leq \mathbb{E}_{\mathbb{P}_{\Gamma}}[h(\beta_p)] + \mathbb{E}_{\mathbb{P}_{\Gamma}^*}[h(\beta_p)] \leq 2M_k(h)$ [see Theorem 6.2 in 47], so $g_{p,t,\mathbf{u}} \in C_p$. Because tV_p is a set of uniqueness of C_p , $g_{p,t,\mathbf{u}}$ is 0 on \mathbb{R} .

Take now $(\mathbf{u}, x) \in U_{p-2} \times \mathbb{R}$ and $g_{p-1,t,\mathbf{u},x} : z \in \mathbb{R} \mapsto \mathcal{F}[\mathbb{P}_{\Gamma} - \mathbb{P}_{\Gamma}^*](t, t\mathbf{u}, z, x) = \mathcal{F}[\mu_{p-1,t,\mathbf{u},x}(\cdot)](z)$ is 0 on tV_{p-1} , where

$$\mu_{p-1,t,\mathbf{u},x}(\cdot) = \int_{\mathbb{R}^p} e^{it(a + \mathbf{b}_{[p-2]}^\top \mathbf{u}_{[p-2]} + \mathbf{b}_p x)} d\mu(a, \mathbf{b}_{[p-2]}, \cdot, \mathbf{b}_p)$$

Because $g_{p-1,t,\mathbf{u},x} \in C_{p-1}$ and tV_{p-1} is set of uniqueness, $g_{p-1,t,\mathbf{u},x}$ is identically 0. We conclude by iterating this argument. \square

Proof of Theorem 2 under $\mathcal{R}_{BC,1}$. Take $0 < \epsilon \leq 1/2$. Assume that \mathbb{P}_{Γ} and \mathbb{P}_{Γ}^* both give rise to the same collection $(\mathbb{E}[\mathbb{1}\{\Gamma^\top \mathbf{s} \geq 0\} | \mathbf{S} = \mathbf{s}])_{\mathbf{s} \in \mathbb{S}_{\mathbf{S}}}$. Recall that, from [24],

$$(A.14) \quad \mathcal{T}f(\mathbf{s}) = \mathcal{T}f^-(\mathbf{s}) = (\mathcal{T}f)^-(\mathbf{s}).$$

One has in $S'(\mathbb{S}^p)$, for all $f \in L^q(\mathbb{S}^p)$,

$$(A.15) \quad \mathcal{T}f = \sum_{m \in \mathbb{N}_0} \lambda_{2m+1,p} Q_{2m+1,p} f,$$

where $\lambda(2m+1, p) = (-1)^m 2\pi^{p/2} 1 \cdot 3 \cdots (2m-1) / (\Gamma(p/2)p(p+2) \cdots (p+2m))$ for all $m \in \mathbb{N}_0$. Hence, we have, for all $\mathbf{x} \in \mathbb{S}_{\mathbf{X}}$,

$$\mathbb{E}[Y | \mathbf{X} = \mathbf{x}] = \sum_{m \in \mathbb{N}_0} \lambda_{2m+1,p} Q_{2m+1,p} f_{\Gamma} \left(\frac{(1, \mathbf{x})}{\sqrt{1 + |\mathbf{x}|_2^2}} \right) + \frac{1}{2}$$

and similarly for f_{Γ}^* . Denote the dot product and Lie norm by $\mathbf{z}^2 = \sum_{k=1}^p \mathbf{z}_k^2$ and $L(\mathbf{z}) = \sqrt{|\mathbf{z}|_2^2 + \sqrt{|\mathbf{z}|_2^4 - |\mathbf{z}^2|^2}}$ respectively. Denote also by $G(\mathbf{z}) = \sum_{m \in \mathbb{N}_0} \lambda_{2m+1,p} Q_{2m+1,p} f_{\Gamma}(\mathbf{z}) + 1/2$ for all $\mathbf{z} \in \mathbb{C}^{p+1}$ (resp. G^* for f_{Γ}^*) and $F(\mathbf{z}) = (f_1(\mathbf{z}), \dots, f_{p+1}(\mathbf{z}))^\top$ for all $\mathbf{z} \in \mathbb{C}^p$, where:

$$f_1(\mathbf{z}) = \frac{1}{\sqrt{1 + \mathbf{z}^2}}, \quad f_2(\mathbf{z}) = \frac{\mathbf{z}_1}{\sqrt{1 + \mathbf{z}^2}}, \quad \dots, \quad f_{p+1}(\mathbf{z}) = \frac{\mathbf{z}_p}{\sqrt{1 + \mathbf{z}^2}}.$$

We have $\mathbb{E}[Y | \mathbf{X} = \mathbf{x}] = G \circ F(\mathbf{x}) = G^* \circ F(\mathbf{x})$ and we prove that $G \circ F, G^* \circ F \in \mathcal{A}(\mathbb{R}^p + i\epsilon \mathbb{B}_{\mathbb{R}}^p)$, hence $(G - G^*) \circ F \in \mathcal{A}(\mathbb{R}^p + i\epsilon \mathbb{B}_{\mathbb{R}}^p)$. For this, we check the conditions of Theorem 1.2.3 in [48].

First, we have $F \in \mathcal{A}(\mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p)$. Indeed, for all $\mathbf{z} \in \mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p$, $1 + \mathbf{z}^2 \in (\mathbb{C} \setminus (-\infty, 0])^{p+1}$. Now, for all $\mathbf{z} \in \mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p$, we have

$$L(F(\mathbf{z}))^2 = \frac{L(1, \mathbf{z})^2}{|\sqrt{1 + \mathbf{z}^2}|^2} \leq \frac{1 + (|\operatorname{Re}(\mathbf{z})|_2 + |\operatorname{Im}(\mathbf{z})|_2)^2}{1 + |\operatorname{Re}(\mathbf{z})|_2^2 - |\operatorname{Im}(\mathbf{z})|_2^2} \leq 1 + 2\epsilon.$$

Moreover, using Lemma 3.23 in [42] for the first display and using the Young inequalities (see [24]) for the third display, we have, for all $\mathbf{z} \in \mathbb{C}^{p+1}$ such that $L(\mathbf{z}) \leq \sqrt{1 + 2\epsilon}$,

$$\begin{aligned} |\lambda_{2m+1,p} Q_{2m+1,p} f_{\Gamma}(\mathbf{z})| &\leq L(\mathbf{z})^{2m+1} \|\lambda_{2m+1,p} Q_{2m+1,p} f_{\Gamma}\|_{L^\infty(\mathbb{S}^p)} \\ &\leq L(\mathbf{z})^{2m+1} \|\mathcal{T}[Q_{2m+1,p} f_{\Gamma}]\|_{L^\infty(\mathbb{S}^p)} \\ &\leq L(\mathbf{z})^{2m+1} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)} \\ &\leq (1 + 2\epsilon)^{m+1/2} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)}. \end{aligned}$$

Now, because $\overline{\lim}_{m \rightarrow \infty} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)}^{1/m} \leq q/(1 + 2\epsilon)$ with $q < 1$, there exists $m_0 \in \mathbb{N}$ such that, for all $m \geq m_0$, $(1 + 2\epsilon)^{m+1/2} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)} \leq \sqrt{1 + 2\epsilon}$. Hence, there exists $C_\epsilon < \infty$ such that, for all $m \in \mathbb{N}_0$, $(1 + 2\epsilon)^{m+1/2} \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)} \leq C_\epsilon$. As a result, for all $\mathbf{z} \in \mathbb{C}^{p+1}$ such that $L(\mathbf{z}) < \sqrt{1 + 2\epsilon}$, we have $\sup_{m \in \mathbb{N}_0} |\lambda_{2m+1,p} Q_{2m+1,p} f_{\Gamma}(\mathbf{z})| < \infty$. Using the fact that $\bar{\mathbf{z}} \mapsto Q_{2m+1,p} f_{\Gamma}(\bar{\mathbf{z}})$ are homogeneous harmonic polynomials and Theorem 1.5.6 in [48] yields $G \circ F \in \mathcal{A}(\mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p)$. Moreover, for all $\mathbf{z} \in \mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p$, we have

$$|G \circ F(\mathbf{z})| \leq \sqrt{1 + 2\epsilon} \sum_{m \in \mathbb{N}_0} \left((1 + 2\epsilon) \|Q_{2m+1,p} f_{\Gamma}\|_{L^1(\mathbb{S}^p)}^{1/m} \right)^m + \frac{1}{2},$$

and the upper bound is a convergent series using $\mathcal{R}_{BC,1}$ (ii). Hence, $G \circ F \in \mathcal{A}(\mathbb{R}^p + i\epsilon\mathbb{B}_{\mathbb{R}}^p)$ is bounded and similarly for $G^* \circ F$, which yields the result. \square

Proof of Theorem 2 under $\mathcal{R}_{BC,2}$. Assume that \mathbb{P}_{Γ} and \mathbb{P}_{Γ}^* both give rise to the same collection $(\mathbb{E}[\Gamma^\top \mathbf{s} \geq 0] | \mathbf{S} = \mathbf{s})_{\mathbf{s} \in \mathbb{S}_{\mathbf{S}}}$. Due to (A.15), \mathcal{T} and Δ commute. For $r \geq 0$, $q \in [1, \infty]$, and $f \in L^q(\mathbb{S}^p)$, $\Delta^k f$ exists in $S'(\mathbb{S}^p)$ and is such that $\Delta^r f = \sum_{m=0}^{\infty} \zeta_{m,p}^r Q_{m,p} f$. When the sum converges in $L^q(\mathbb{S}^p)$, f is said to belong to $W_q^{2r}(\mathbb{S}^p)$. This Sobolev space is equipped with $\|f\|_{W_q^{2r}(\mathbb{S}^p)} = \|f\|_{L^q(\mathbb{S}^p)} + \|\Delta^r f\|_{L^q(\mathbb{S}^p)}$. Using that Δ commutes with \mathcal{T} , (A.14), the Young inequalities (see [24]), and $\mathcal{R}_{BC,2}$ (ii), yield that, for all $m \in \mathbb{N}_0$, $\mathcal{T}f \in W_\infty^m(\mathbb{S}^p)$ and

$$(A.16) \quad \|\Delta^m \mathcal{T}f\|_{L^\infty(\mathbb{S}^p)} \leq \|\Delta^m f^-\|_{L^1(\mathbb{S}^p)}.$$

Denote, for $T \in \mathbb{N}$, by $\mathcal{K}_T f(\cdot) = \int_{\mathbb{S}^p} k_T(\cdot, \mathbf{s}) f(\mathbf{s}) d\sigma(\mathbf{s})$, where $k_T(\mathbf{x}, \mathbf{y}) = \sum_{l=0}^T \psi(l/T) q_{l,p}(\mathbf{x}, \mathbf{y})$ with $\psi \in C^\infty([0, \infty))$, nonnegative, nonincreasing, such that $\psi(x) = 1$ if $x \in [0, 1]$, $0 \leq \psi(x) \leq 1$ if $x \in [1, 2]$, and $\psi(x) = 0$ if $x \geq 2$, Δ also commutes with \mathcal{K}_T . Now,

using that Δ commutes with \mathcal{T} and \mathcal{K}_T (first and third displays) and the Young inequalities (first display), we obtain, for all $m \in \mathbb{N}_0$ and $k \in \mathbb{N}$,

$$\begin{aligned}
 \|\Delta^m \mathcal{T}f - \mathcal{T}\mathcal{K}_T \Delta^m f\|_\infty &\leq \|\Delta^m f^- - \Delta^m \mathcal{K}_T f^-\|_{L^1(\mathbb{S}^p)} \\
 &= \|\Delta^m f^- - \mathcal{K}_T \Delta^m f^-\|_{L^1(\mathbb{S}^p)} \\
 &\leq CT^{-2k} \|\Delta^m f^-\|_{W_1^{2k}(\mathbb{S}^p)} \\
 (A.17) \qquad &\leq CT^{-2k} \left(\|f^-\|_{W_1^{2m}(\mathbb{S}^p)} + \|f^-\|_{W_1^{2m+2k}(\mathbb{S}^p)} \right),
 \end{aligned}$$

where the third display follows from Proposition A.2 in [24]. Using

$$\mathcal{T}\mathcal{K}_T \Delta^m f = \sum_{l=0}^{\lfloor (T-1)/2 \rfloor} \psi\left(\frac{2l+1}{T}\right) \lambda_{2l+1,p} Q_{2l+1,p} \Delta^m f,$$

that $q_{2l+1,p}(x, y)$ is a polynomial in $\mathbf{x}^\top \mathbf{y}$, and that $\Delta^m f \in L^1(\mathbb{S}^p)$, we obtain that $\mathcal{T}\mathcal{K}_T \Delta^m f$ is odd and continuous and conclude from (A.17) and $\mathcal{R}_{BC,2}$ (ii) that $\mathcal{T}f \in C_{\text{odd}}^\infty(\mathbb{S}^p)$. $\mathcal{R}_{BC,2}$ (ii) and (A.16) yield

$$\|\Delta^m \mathcal{T}(f_\Gamma - f_\Gamma^*)\|_{L^\infty(\mathbb{S}^p)} \leq \widetilde{M}_m,$$

where $\widetilde{M}_m = \max(c_{f_\Gamma} + c_{f_\Gamma^*}, 1) \max(b_{f_\Gamma}, b_{f_\Gamma^*})^m M_m$ is such that

$$\sum_{m \in \mathbb{N}} \widetilde{M}_m^{1/m} \geq \max(c_{f_\Gamma} + c_{f_\Gamma^*}, 1) \max(b_{f_\Gamma}, b_{f_\Gamma^*}) \sum_{m \in \mathbb{N}} M_m^{1/m} = \infty$$

because $\{M_m\}$ satisfies (12). Also, (15), the fact that $\mathcal{T}(f_\Gamma - f_\Gamma^*) \in C_{\text{odd}}^\infty(\mathbb{S}^p)$, and $\mathcal{R}_{BC,2}$ (i) imply that, for all $\mathbf{u} \in U$ and $m \in \mathbb{N}_0$, $\Delta^m \mathcal{T}(f_\Gamma - f_\Gamma^*)(\mathbf{u}) = 0$. From the proof of Theorem 10 in [8] (paragraph after Lemma 5) we obtain, for all $\mathbf{u} \in U$ and $l \in \mathbb{N}_0$, $Q_{2l+1,p} \mathcal{T}(f_\Gamma - f_\Gamma^*)(\mathbf{u}) = 0$. Hence, by $\mathcal{R}_{BC,2}$ (i) and the fact that $Q_{2l+1,p} \mathcal{T}(f_\Gamma - f_\Gamma^*)$ is the restriction to \mathbb{S}^p of a harmonic homogeneous polynomial of degree $2l+1$, we obtain, for all $\mathbf{s} \in \mathbb{S}^p$ and $l \in \mathbb{N}_0$, $Q_{2l+1,p} \mathcal{T}(f_\Gamma - f_\Gamma^*)(\mathbf{s}) = 0$. Thus $(f_\Gamma - f_\Gamma^*)^- = 0$ and the odd parts of f_Γ and f_Γ^* coincide in $L^1(\mathbb{S}^p)$.

By the argument in [25], using that for $f = f_\Gamma$ or $f = f_\Gamma^*$, for a.e. $\mathbf{u} \in \mathbb{S}^p$ $f(\mathbf{u}) \geq 0$, the definition of f^- , and $f_\Gamma(\mathbf{u})f_\Gamma(-\mathbf{u}) = 0$ for a.e. $\mathbf{u} \in \mathbb{S}^p$, we obtain that, for a.e. $\mathbf{u} \in \mathbb{S}^p$, $f(\mathbf{u}) = 2f^-(\mathbf{u})\mathbb{1}\{f^-(\mathbf{u}) > 0\}$. Thus, there is a one-to-one mapping between f and f^- . This allows to conclude. \square

Proof of Theorem 3. Assume that $P = \mathbb{P}_\Gamma$ and $P = \mathbb{P}_\Gamma^*$ both give rise to the same pair $(\mathbb{P}_{\Gamma_1+\Gamma_3}, \mathbb{P}_{\Gamma_2+\Gamma_3})$. For $t_0 > 0$ small enough, on $(-t_0, t_0)$, $\varphi_\delta, \varphi_\delta^*, \varphi_{\epsilon_1}, \varphi_{\epsilon_1}^*, \varphi_{\epsilon_2}, \varphi_{\epsilon_2}^*$ do not vanish, hence there exist nonvanishing continuous functions p_δ, p_{ϵ_1} , and p_{ϵ_2} such that $\varphi_\delta = p_\delta \varphi_\delta^*, \varphi_{\epsilon_1} = p_{\epsilon_1} \varphi_{\epsilon_1}^*, \varphi_{\epsilon_2} = p_{\epsilon_2} \varphi_{\epsilon_2}^*$. By the restriction \mathcal{R}_K , we have $p_{\epsilon_1} \in C^\infty(-t_0, t_0)$. We have

$$(A.18) \qquad p_\delta(\mathbf{t}_1 + \mathbf{t}_2) p_{\epsilon_1}(\mathbf{t}_1) p_{\epsilon_2}(\mathbf{t}_2) = 1 \quad \text{for all } -t_0 < \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_1 + \mathbf{t}_2 < t_0,$$

hence

$$(A.19) \qquad p_\delta(t) p_{\epsilon_1}(t) = 1 \quad \text{and} \quad p_\delta(t) p_{\epsilon_2}(t) = 1 \quad \text{for all } -t_0 < t < t_0.$$

Injecting (A.19) into (A.18) we obtain, for all $-t_0 < \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_1 + \mathbf{t}_2 < t_0$, $p_\delta(\mathbf{t}_1 + \mathbf{t}_2) = p_\delta(\mathbf{t}_1)p_\delta(\mathbf{t}_2)$, which, using again (A.19), yields $p_{\epsilon_1}(\mathbf{t}_1 + \mathbf{t}_2) = p_{\epsilon_1}(\mathbf{t}_1)p_{\epsilon_1}(\mathbf{t}_2)$. Hence, we have $p'_{\epsilon_1}(\mathbf{t}_1 + \mathbf{t}_2) = p'_{\epsilon_1}(\mathbf{t}_1)p_{\epsilon_1}(\mathbf{t}_2)$, which at $\mathbf{t}_1 = 0$ and $\mathbf{t}_2 = t$ yields

$$(A.20) \quad p'_{\epsilon_1}(t) = p'_{\epsilon_1}(0)p_{\epsilon_1}(t) \quad \text{for all } -t_0 < t < t_0,$$

where $p'_{\epsilon_1}(0) = \varphi'_{\epsilon_1}(0) - (\varphi^*_{\epsilon_1})'(0) = i(\mathbb{E}_{\mathbb{P}_{\epsilon_1}}[\epsilon_1] - \mathbb{E}_{\mathbb{P}^*_{\epsilon_1}}[\epsilon_1]) \in i\mathbb{R}$ which we denote by $p'_{\epsilon_1}(0) := ib$. Thus, we obtain, for all $t_0 < t < t_0$, $p_{\epsilon_1}(t) = \exp(ib t)$. This yields that $t \mapsto \varphi_{\epsilon_1}(t) - \exp(ib t)\varphi^*_{\epsilon_1}(t)$ is 0 on $(-t_0, t_0)$. Thus, using \mathcal{R}_K , we have $\varphi_{\epsilon_1}^{(1)}(0) = ib + (\varphi^*_{\epsilon_1})^{(1)}(0)$ then using that \mathbb{P}_{ϵ_1} and $\mathbb{P}^*_{\epsilon_1}$ are both mean 0 yields $b = 0$. Thus, \mathbb{P}_{ϵ_1} and $\mathbb{P}^*_{\epsilon_1}$ have the same moments so $\mathbb{P}_{\epsilon_1} = \mathbb{P}^*_{\epsilon_1}$ by \mathcal{R}_K .

Now, for all $t \in \mathbb{R}$, $\varphi_{\mathbf{Y}_1}(t) = \varphi_\delta(t)\varphi^*_{\epsilon_1}(t) = \varphi^*_\delta(t)\varphi^*_{\epsilon_1}(t)$, hence, because the zeros of $\varphi^*_{\epsilon_1}$ are isolated [see Lemma 4.8 in 1] and φ_δ and φ^*_δ are continuous, we obtain $\varphi_\delta = \varphi^*_\delta$. Similarly, because, for all $t \in \mathbb{R}$, $\varphi_{\mathbf{Y}_2 - \mathbf{Y}_1}(t) = \varphi^*_{\epsilon_2}(t)\varphi^*_{\epsilon_1}(-t) = \varphi_{\epsilon_2}(t)\varphi^*_{\epsilon_1}(-t)$, the zeros of $\varphi^*_{\epsilon_1}$ are isolated and φ_{ϵ_2} and $\varphi^*_{\epsilon_2}$ are continuous, we obtain $\varphi_{\epsilon_2} = \varphi^*_{\epsilon_2}$, hence the result. \square

Proof of Theorem 4. We use $\mathcal{X}_0 = \{\mathbf{x} \in \mathbb{R}^T : \prod_{j=1}^T \mathbf{x}_j \neq 0, \prod_{m \neq j} (\mathbf{x}_m - \mathbf{x}_j) \neq 0\}$ and, for all $(\mathbf{v}, \mathbf{x}) \in \mathbb{R}^T \times \mathcal{X}_0$,

$$(A.21) \quad \Theta : (\mathbf{v}, \mathbf{x}) \mapsto \sum_{k=1}^T \mathbf{v}_k \left(\sum_{j=1}^T b_{jk}(\mathbf{x}) \mathbf{x}_j^T \right),$$

$$\left\{ \begin{array}{l} b_{jk}(\mathbf{x}) = \frac{(-1)^{k+1}}{\prod_{\substack{1 \leq m \leq T \\ m \neq j}} (\mathbf{x}_m - \mathbf{x}_j)} \sum_{\substack{1 \leq i_1 < \dots < i_{T-k} \leq T \\ i_1, \dots, i_{T-k} \neq j}} \mathbf{x}_{i_1} \mathbf{x}_{i_2} \dots \mathbf{x}_{i_{T-k}} \quad \text{for all } k \neq T \\ b_{jk}(\mathbf{x}) = \frac{(-1)^{T+1}}{\prod_{\substack{1 \leq m \leq T \\ m \neq j}} (\mathbf{x}_m - \mathbf{x}_j)} \quad \text{for } k = T. \end{array} \right.$$

Assume that \mathbb{P}_Γ and \mathbb{P}^*_Γ both give rise to the same collections $\left(\mathbb{P}_{\Gamma_1 + \sum_{j=1}^T \Gamma_{j+1} \mathbf{x}_t^j + \epsilon_t} \right)_{\mathbf{x}_t \in \mathbb{S}_{\mathbf{X}_t}, t=1, \dots, T}$. Using $\mathcal{R}_{LP,0}$ (ii), there exists $r > 0$ such that $r\mathbf{1}_c \in \mathcal{X}_1$ and conditioning on $\mathbf{X} = r\mathbf{1}_c$ and using $\delta = \alpha + \sum_{k=1}^T \beta_k r^k$ yield, for all $\mathbf{t} \in \mathbb{R}^T$, $\mathbb{E} \left[e^{i\mathbf{t}^\top \mathbf{Y}} \mid \mathbf{X} = r\mathbf{1}_c \right] = \varphi_\delta \left(\sum_{j=1}^T \mathbf{t}_j \right) \prod_{j=1}^T \varphi_{\epsilon_j}(\mathbf{t}_j)$, hence

$$\varphi_\delta \left(\sum_{j=1}^T \mathbf{t}_j \right) \prod_{j=1}^T \varphi_{\epsilon_j}(\mathbf{t}_j) = \varphi^*_\delta \left(\sum_{j=1}^T \mathbf{t}_j \right) \prod_{j=1}^T \varphi^*_{\epsilon_j}(\mathbf{t}_j).$$

Then, following the same steps as in the proof of Theorem 3 yields $\mathbb{P}_{\epsilon_1} = \mathbb{P}^*_{\epsilon_1}$. Now, for all $t \in \mathbb{R}$,

$$\varphi_\delta(t)\varphi^*_{\epsilon_1}(t) = \varphi^*_\delta(t)\varphi^*_{\epsilon_1}(t),$$

hence, because the zeros of $\varphi^*_{\epsilon_1}$ are isolated [see Lemma 4.8 in 1] and φ_δ and φ^*_δ are continuous, we obtain $\varphi_\delta = \varphi^*_\delta$. Similarly, because, for all $t \in \mathbb{R}$ and $j = 2, \dots, T$, $\mathbb{E} \left[e^{it(\mathbf{Y}_j - \mathbf{Y}_1)} \mid (\mathbf{X}_1, \mathbf{X}_j) = (r, r) \right] = \varphi^*_{\epsilon_j}(t)\varphi^*_{\epsilon_1}(-t)$, we have

$$\varphi^*_{\epsilon_j}(t)\varphi^*_{\epsilon_1}(-t) = \varphi_{\epsilon_j}(t)\varphi^*_{\epsilon_1}(-t).$$

Thus, using that the zeros of $\varphi_{\epsilon_1}^*$ are isolated and φ_{ϵ_j} and $\varphi_{\epsilon_j}^*$ are continuous for all $j = 2, \dots, T$, we obtain $\varphi_{\epsilon_j} = \varphi_{\epsilon_j}^*$ for all $j = 2, \dots, T$, hence, for all $\mathbf{t} \in \mathbb{R}^T$ and $\mathbf{x} \in \mathbb{S}_{\mathbf{X}}$,

$$(A.22) \quad \varphi_{\alpha, \beta} \left(\sum_{j=1}^T \mathbf{t}_j, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j, \dots, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j^T \right) = \varphi_{\alpha, \beta}^* \left(\sum_{j=1}^T \mathbf{t}_j, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j, \dots, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j^T \right).$$

Then, for all $\mathbf{x} \in \mathcal{X}_0$, we use a change of variable that relates $\mathbf{t} \in \mathbb{R}^T$ to $\mathbf{v} \in \mathbb{R}^T$ such that $(\sum_{j=1}^T \mathbf{t}_j, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j, \dots, \sum_{j=1}^{T-1} \mathbf{t}_j \mathbf{x}_j^{T-1}) = \mathbf{v}$. This change of variable allows to choose the values of the $T - 1$ first variables in (A.22) independently from each other. This can be written as $\mathbf{t} = (V^{-1}(\mathbf{x}))^\top \mathbf{v}$, where

$$V(\mathbf{x}) = \begin{pmatrix} 1 & \mathbf{x}_1 & \mathbf{x}_1^2 & \dots & \mathbf{x}_1^{T-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \mathbf{x}_T & \mathbf{x}_T^2 & \dots & \mathbf{x}_T^{T-1} \end{pmatrix}$$

is the Vandermonde matrix. We use $D(\mathbf{x})$, the diagonal matrix which entries are the coordinates of \mathbf{x} . It is a classical result that, for all $\mathbf{x} \in \mathcal{X}_0$, $D(\mathbf{x})$ and $V(\mathbf{x})$ are invertible hence $\tilde{V}(\mathbf{x}) = (D(\mathbf{x})V(\mathbf{x}))^\top$ is invertible. Then, for all $\mathbf{v} \in \mathbb{R}^T$ and $\mathbf{x} \in \mathcal{X}_0$, we can express \mathbf{t} as a function of \mathbf{v} and \mathbf{x} , and obtain for the last variable in

$$\varphi_{\alpha, \beta}^* \left(\sum_{j=1}^T \mathbf{t}_j, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j, \dots, \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j^T \right),$$

that

$$(A.23) \quad \sum_{j=1}^T \mathbf{t}_j \mathbf{x}_j^T = \mathbf{t}^\top \mathbf{x}^T = \left((V(\mathbf{x})^{-1})^\top \mathbf{v} \right)^\top \mathbf{x}^T = \left(D(\mathbf{x}) \tilde{V}(\mathbf{x})^{-1} \mathbf{v} \right)^\top \mathbf{x}^T = \Theta(\mathbf{v}, \mathbf{x}).$$

Using (A.22), this yields, for all $\mathbf{v} \in \mathbb{R}^T$ and $\mathbf{x} \in \mathcal{X}_0$,

$$(A.24) \quad \varphi_{\alpha, \beta}(\mathbf{v}, \Theta(\mathbf{v}, \mathbf{x})) = \varphi_{\alpha, \beta}^*(\mathbf{v}, \Theta(\mathbf{v}, \mathbf{x})).$$

Then, using that the vector space spanned by $\mathcal{F}[\mathcal{P}_c(\Omega_T)]$ is a quasi-analytic class of functions on \mathbb{R} , we obtain that, for all $\mathbf{v} \in \mathbb{R}^T$, $z \in \mathbb{R} \mapsto (\varphi_{\alpha, \beta} - \varphi_{\alpha, \beta}^*)(\mathbf{v}, z)$ belongs to a quasi-analytic class. Finally, for all $\mathbf{v} \in \mathbb{R}^T \setminus \mathcal{V}$ where \mathcal{V} is such that $\mathbb{R}^T \setminus \mathcal{V}$ is dense in \mathbb{R}^T , \mathcal{R}_{LP} (i) and that $\mathbf{x} \mapsto \Theta(\mathbf{v}, \mathbf{x})$ is continuous on $\mathbb{S}_{\mathbf{X}} \cap \mathcal{X}_0$ yield that $U_{T, \mathbf{v}} = \{\Theta(\mathbf{v}, \mathbf{u}), \forall \mathbf{u} \in \mathbb{S}_{\mathbf{X}} \cap \mathcal{X}_0\}$ contains a bounded sequence of distinct points. We obtain that, for all $\mathbf{v} \in \mathbb{R}^T \setminus \mathcal{V}$ and $z \in \mathbb{R}$, $\varphi_{\alpha, \beta}(\mathbf{v}, z) = \varphi_{\alpha, \beta}^*(\mathbf{v}, z)$ hence $\mathbb{P}_{\alpha, \beta} = \mathbb{P}_{\alpha, \beta}^*$ by continuity for all $\mathbf{v} \in \mathbb{R}^T$. \square

Proof of Theorem 5. Let $0 < \epsilon < 1$. Assume that \mathbb{P}_{Γ} and \mathbb{P}_{Γ}^* both give rise to the same collections $\left(\mathbb{P}_{f(\Gamma^\top \mathbf{x}_t) + \eta_t} \right)_{\mathbf{x}_t \in \mathbb{S}_{\mathbf{X}_t}, t=1,2}$. Denote by $\mathcal{X}_1 = \{(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{S}_{\mathbf{X}_1, \mathbf{X}_2} : \mathbf{x}_1 = \mathbf{x}_2\}$.

Using \mathcal{R}_{SI} (i), there exists $r > 0$ such that $r(\mathbf{1}_c, \mathbf{1}_c) \in \mathcal{X}_1$, hence, for all $\mathbf{t} \in \mathbb{R}^2$, using $\delta := f(\Gamma^\top(r\mathbf{1}_c))$, we have

$$(A.25) \quad \varphi_\delta(\mathbf{t}_1 + \mathbf{t}_2) \varphi_{\eta_1}(\mathbf{t}_1) \varphi_{\eta_2}(\mathbf{t}_2) = \varphi_\delta^*(\mathbf{t}_1 + \mathbf{t}_2) \varphi_{\eta_1}^*(\mathbf{t}_1) \varphi_{\eta_2}^*(\mathbf{t}_2).$$

Then, following the same steps as in the proof of Theorem 3 yields $\varphi_{\eta_1} = \varphi_{\eta_1}^*$. Now, for all $t \in \mathbb{R}$, $\varphi_{\mathbf{Y}_1 | \mathbf{X}_1}(t | r\mathbf{1}_c) = \varphi_\delta(t) \varphi_{\eta_1}^*(t) = \varphi_\delta^*(t) \varphi_{\eta_1}^*(t)$, hence, because the zeros of $\varphi_{\eta_1}^*$ are

isolated [see Lemma 4.8 in 1] and φ_δ and φ_δ^* are continuous, we obtain $\varphi_\delta = \varphi_\delta^*$. Similarly, because, for all $t \in \mathbb{R}$, $\varphi_{\mathbf{Y}_2 - \mathbf{Y}_1 | \mathbf{X}_1, \mathbf{X}_2}(t | (r\mathbf{1}_c, r\mathbf{1}_c)) = \varphi_{\boldsymbol{\eta}_2}^*(t)\varphi_{\boldsymbol{\eta}_1}^*(-t) = \varphi_{\boldsymbol{\eta}_2}(t)\varphi_{\boldsymbol{\eta}_1}^*(-t)$, the zeros of $\varphi_{\boldsymbol{\eta}_1}^*$ are isolated and $\varphi_{\boldsymbol{\eta}_2}$ and $\varphi_{\boldsymbol{\eta}_2}^*$ are continuous, we obtain $\varphi_{\boldsymbol{\eta}_2} = \varphi_{\boldsymbol{\eta}_2}^*$.

Thus, we obtain, for all $\mathbf{t} \in \mathbb{R}^2$ and $(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{S}_{\mathbf{X}_1, \mathbf{X}_2}$,

$$(A.26) \quad \varphi_{f(\boldsymbol{\Gamma}^\top \mathbf{x}_1), f(\boldsymbol{\Gamma}^\top \mathbf{x}_2)}(\mathbf{t}) = \varphi_{f(\boldsymbol{\Gamma}^\top \mathbf{x}_1), f(\boldsymbol{\Gamma}^\top \mathbf{x}_2)}^*(\mathbf{t}).$$

This amounts to study identification in $\mathbf{Z}_t = f(\boldsymbol{\Gamma}^\top \mathbf{X}_t)$, $t = 1, 2$ and $\mathbf{Z}_2 \geq \mathbf{Z}_1$ is equivalent to $\boldsymbol{\Gamma}^\top (\mathbf{X}_2 - \mathbf{X}_1) \geq 0$. This is the binary choice model, hence Theorem 2 under $\mathcal{R}_{BC,2}$ yields the result. \square