A class of random utility models yielding the exploded logit

Karim Kilani

A Class of Random Utility Models Yielding the Exploded Logit

Karim Kilani[a,*]

[a]*LIRSA, Conservatoire National des Arts et Métiers, Paris, 75003, France*

Abstract

We reexamine a family of distributions introduced within the framework of random utility models by David Strauss. This family generates ranking probabilities of the exploded logit model and, de facto, the choice probabilities of the multinomial logit model. We explore the necessary and sufficient conditions for its validity within the copula theory. By specifying the minimal assumptions required for the support of the marginal utility distributions, we clarify and reinforce the fundamental structure of the model, proving that it relies on strict archimedean copulas. Additionally, we provide a new mathematical proof by induction on the number of alternatives confirming that these utility distributions indeed generate the exploded logit model.

Keywords: Archimedean copulas, Choice probabilities, Copula theory, Exploded logit model, Multinomial logit model, Ranking probabilities, Random utility models

1. Introduction

In a seminal publication more than four decades ago, Strauss (1979) presented groundbreaking findings. One of his key contributions was the expansion of the multivariate distribution of utilities, moving beyond inde-

*Corresponding author
Email address: karkil2205@gmail.com (Karim Kilani)

pendent double-exponentials (or Gumbels) to generate choice probabilities that are consistent with the multinomial logit model. He introduced a significant generalization, referred to in this paper as the *Strauss model*, which incorporates correlated double-exponentials and extends to a broader class of distributions. This paper will examine and define this generalization in detail.

Although the Strauss model introduces a specific type of correlation among the alternatives and perturbs the original double-exponential marginals, the choice probabilities remain invariant and align with the logit formula, which Strauss refers to as the "choice axiom", remaining faithful to the terminology of the foundational work by Luce (1959).

At a more detailed level of choice, the ranking probabilities—i.e., the probabilities associated with all possible rankings or permutations of the available alternatives—follow the exploded logit model, which Strauss designates as "complete decomposition". It is crucial to note that if the ranking probabilities are consistent with the exploded logit model, then the choice probabilities must necessarily follow the multinomial logit model. However, whether the converse is true—are there random utility models for which choice probabilities are consistent with the multinomial logit model, but ranking probabilities do not follow the exploded logit model?—remains an unresolved question.[1]

In this paper, we revisit the Strauss model by presenting minimal assumptions regarding the support of the marginal distributions, followed by a discussion of the necessary and sufficient conditions for the model's validity, drawing upon contemporary copula theory (for an authoritative reference, see Nelsen, 2006). Although Strauss did not employ copulas in his work, we prove that his model aligns with archimedean copulas that are applied

[1]Initially, we became interested in the Strauss model while seeking an answer to this question. In examining this model closely, we decided to write this paper to detail it further and hope to address this open question in the future.

to specific marginals. We draw upon the existing literature on archimedean copulas to establish the conditions for the model's validity. This methodology capitalizes on the advanced theoretical constructs offered by archimedean copulas, thus providing rigor and precision to our findings by relying on well-established results. Consequently, we circumvent the redundancy of complex proofs, focusing on elucidating the essential conditions while ensuring a robust and precise understanding of the model.

It is noteworthy that copulas, particularly archimedean copulas, have been studied within the discrete choice models framework. Schwiebert (2016) proves that modeling the multivariate distribution of utilities with archimedean copulas aids the computation of integrals necessary for determining choice probabilities. This approach yields closed-form expressions for the partial derivatives of the cumulative distribution functions, which act as integrands within these computations. Consequently, the integration of probabilities is simplified through the use of techniques including Gauss-Hermite quadrature or Monte Carlo methods. It is not our objective to furnish an exhaustive survey of the literature on the application of copulas to random utility models; rather, our emphasis lies exclusively on utilizing copulas to define the boundaries of the family of distributions encompassed by the Strauss model.

It should be noted that Beggs et al. (1981) provided an elegant proof, although two years after Strauss (1979), establishing that the multinomial logit model, derived from a random utility model with independent double-exponentials, results in the exploded logit model. However, Strauss showed that the exploded logit model can be achieved within a much broader framework that includes a larger family of distributions. By revisiting the Strauss model, we highlight the broader scope of his results, which include many models that conform to the exploded logit model beyond the specific case of independent double-exponentials.

The subsequent sections of this paper are structured as follows. Section 2

elaborates on the formulation of the Strauss model alongside the underlying assumptions. Section 3 provides a representation of the Strauss model via copulas, elucidating that this representation utilizes archimedean copulas. Section 4 details examples of distributions that conform to the framework of the Strauss model. In Section 5, we explicate the conformity of the ranking probabilities within the Strauss model to the exploded logit model, providing a new proof of this claim. Lastly, Section 6 summarizes the findings of the paper and suggests avenues for future research.

2. Strauss model: formulation and assumptions

Consider a random utility model where an individual's choices are probabilistically described by a vector of random utilities $\mathbf{U} = (U_1, \ldots, U_n)$, with each U_i representing the utility of alternative i among n possible alternatives (i ranging from 1 to n and $n \geq 2$). Our analysis focuses on the scenario where this vector follows a distribution introduced by Strauss (1979), which characterizes the Strauss model. The cumulative distribution function (CDF) of the utility vector $\mathbf{U}$ is given by:

$$F_{\boldsymbol{U}}(x_1, \ldots, x_n) = \psi\left(\sum_{i=1}^{n} e^{\frac{v_i - x_i}{\sigma}}\right), \tag{1}$$

where x_i is a realization of U_i, belonging to $(-\infty, +\infty)$, and ψ is a function whose properties will be discussed in detail later.

We chose to present the arguments in the form $(v_i - x_i)/\sigma$ because this convention is widely used in the representation of families of probability distributions relative to a standardized reference distribution. Here, v_i denotes the deterministic component of the utility of alternative i, which serves as a location parameter, while σ is a strictly positive scale parameter.[2] This

[2] Although it is possible to generalize the model by introducing distinct scale parameters for each alternative, such a generalization would compromise the exploded logit model.

formulation clarifies how the utilities depend on these parameters. Despite a slight modification from the original formulation, this representation remains consistent with the Strauss model and preserves its foundational structure.

In order to streamline our notation and facilitate subsequent computations, we reformulate the utility vector $\mathbf{U}$ into an alternative vector $\mathbf{T} = (T_1, \ldots, T_n)$, which is conceptualized as a vector of disutilities, via the transformation:

$$\mathbf{T} = \exp\left(-\frac{\mathbf{U}}{\sigma}\right).$$

Conversely, the original utilities can be restored using the relationship:

$$\mathbf{U} = -\sigma \ln \mathbf{T}.$$

Therefore, the survival function (SF) corresponding to the transformed vector $\mathbf{T}$ can be written in a slightly more compact form than the CDF of $\mathbf{U}$ as:

$$\bar{F}_{\mathbf{T}}(t_1, \cdots, t_n) = \psi\left(\sum_{i=1}^{n} \omega_i t_i\right), \tag{2}$$

where $t_i \in [0, +\infty)$. Here, $\omega_i \equiv e^{v_i/\sigma}$ denotes a transformed location parameter for each alternative i. The vector $\mathbf{T}$ is parameterized by $\omega_1, \ldots, \omega_n$, and the scale parameter σ. Thus, SF is expressed in terms of $\boldsymbol{\omega}$ and σ. This formulation will be referred to as the *transformed Strauss model*, and Eq. (2) will be used to define the SF of $\mathbf{T}$.

Strauss (1979) outlined certain necessary conditions to establish the validity of the CDF defined by Eq. (1). However, these conditions are not exhaustive. Although additional conditions were introduced incrementally throughout his derivations, they were not explicitly stated at the outset of his paper. In this analysis, we will present the necessary and sufficient conditions to validate the survival function (SF) $\mathbf{T}$, as defined by Eq. (2), and, consequently, to validate the utility distribution function $\mathbf{U}$, as defined by Eq. (1).

Validating a multivariate CDF or SF is a complex task. It requires satisfying rectangle inequalities, which involve intricate alternating sums to ensure that the probabilities computed for hyperrectangles are well-defined and non-negative (for a detailed discussion of these conditions, the reader is referred to Joe, 1997, pp. 11-12).

The distribution of the utility vector within the Strauss model $\mathbf{U}$, or conversely, the transformed Strauss model via the disutility vector $\mathbf{T}$, necessitates the fulfilment of specific conditions to guarantee the validity of the associated multivariate CDF or SF. To guarantee this, it is imperative that the function ψ satisfies the specified conditions, enabling the multivariate functions defined through ψ to meet the requisite properties in hyperrectangles. These properties, which encompass monotonicity, continuity, boundedness, and non-negativity in hyperrectangles, are essential to affirm the validity of a multivariate CDF or SF.

As an initial phase in the examination of these models, it is imperative to determine the marginal distributions of the variables U_i and T_i. The derivation of these distributions is achievable by examining the asymptotic behavior of the multivariate distributions as the realizations of the other variables approach infinity.

The marginal CDF of U_i is derived considering the asymptotic limit of the multivariate CDF $F_{\mathbf{U}}$, as given in Eq. (1), when the realizations of variables other than x_i approach infinity. Consequently, the marginal CDF for U_i is:

$$F_{U_i}(x_i) = \psi\left(e^{\frac{v_i - x_i}{\sigma}}\right), \quad i = 1, \ldots, n. \tag{3}$$

In parallel, the marginal SF for T_i, which is the transform of U_i with CDF given in Eq. (3), is obtained by taking the asymptotic limit of the multivariate SF $\bar{F}_{\mathbf{T}}$, as indicated in Eq. (2), when the realizations other than t_i tend to infinity. The marginal SF for T_i, which is the counterpart of

the CDF given in Eq. (3), is then:

$$\bar{F}_{T_i}(t_i) = \psi\left(\omega_i t_i\right), \quad i = 1, \ldots, n. \tag{4}$$

The aforementioned equation proves that ψ acts as a univariate SF within the interval $[0, \infty)$. To facilitate a more straightforward specification of marginal distributions, ψ is selected with an unbounded support that encompasses zero, which extends throughout the positive half-line. This choice mitigates the analytical complications associated with more constrained supports. Consequently, by opting for this support, all T_i share the same support, $[0, \infty)$, regardless of their location parameters. This simplifies the analysis and precludes potential complications that may arise from alternative support types, such as corner difficulties. Moreover, this assumption implies that, should we revert to the original utility U_i, its support would extend across the entire real line.[3] These initial observations culminate in the ensuing assumption.

Assumption 1. *The function ψ satisfies the following conditions:*

- *It is strictly decreasing and continuous over $[0, \infty)$.*
- $\psi(0) = 1$ *and* $\psi(+\infty) \equiv \lim_{x \to +\infty} \psi(x) = 0$.

In calculating the densities of the utility subvectors (or disutilities), the successive derivatives of the function ψ are crucial. For example, to derive the marginal density of the variable U_1, the first derivative of ψ is needed.

[3] If ψ had a bounded support $[a, b]$ with $0 < a < b < \infty$, the support of T_i would become $[a/\omega_i, b/\omega_i]$. As ω_i becomes small, the support of T_i extends to the right, which decreases the probability that T_i will achieve the minimum and may even result in the complete elimination of that alternative from consideration. To avoid this problem and ensure that all alternatives have a non-zero probability of being selected, we chose an unbounded support that includes zero, $[0, \infty)$. Although we could have chosen a support of the type $[0, b)$ with b finite, this choice would have introduced unnecessary complications, such as corner difficulties, without adding a significant benefit to the analysis.

This derivative must be strictly negative, as the support of ψ is chosen over the interval $[0, \infty)$ to ensure that the probability density is strictly positive, in accordance with Assumption 1.

To obtain the joint density of the pair (U_1, U_2), the second derivative of ψ is used, and this pattern continues for higher-dimensional subvectors. As more components are added to a subvector, the required order of the derivative of ψ increases. Consequently, each subvector density corresponds to a specific order of derivative of ψ. The signs of these successive derivatives are essential to ensure that the marginal densities and subvector densities are correctly defined and meet the necessary sign requirements.

Therefore, an additional assumption is needed to guarantee the model's validity. Specifically, the following assumption is made to ensure that the marginal probability densities meet the required sign conditions.

Assumption 2. *The function ψ is completely monotonic on the interval $[0, \infty)$. In other terms, ψ is infinitely differentiable on $[0, \infty)$, and for every integer $k \geq 0$, its successive derivatives satisfy:*

$$(-1)^k \psi^{(k)}(y) \geq 0, \quad \forall y \in [0, \infty). \tag{5}$$

The complete monotonicity of ψ necessitates infinite differentiability. The specified signs guarantee that the marginal probability densities observed when selecting a particular number of utilities or disutilities fulfill the requisite conditions, thus enabling the model to support an arbitrary number of alternatives. In the subsequent section, we will employ copula theory to illustrate that Assumptions 1 and 2 are both necessary and sufficient to establish the model's validity, particularly in the context of complete supports and accommodating an arbitrary number of alternatives. Although these assumptions could be moderated for constrained numbers of alternatives, we elect to retain these general conditions to preserve the model's flexibility.

These assumptions are essential for ensuring the validity of the Strauss

model and its transformed counterpart, and their sufficiency will be proved through the application of copula theory, a rigorous and well-established framework. Rather than redeveloping existing mathematical derivations, we will utilize established results from copula theory to substantiate the sufficiency of the proposed assumptions. The subsequent section will elucidate the role of copula theory in confirming the sufficiency of our assumptions, drawing upon the robust findings within this domain.

3. Representation by archimedean copulas

For the subsequent analysis, we choose to analyze the distribution of the disutility vector $\mathbf{T}$ instead of the utility vector $\mathbf{U}$, as it is more advantageous. This preference enables a more rigorous and simplified representation of the Strauss model via the employment of a copula-based methodology.

A pivotal theorem within the domain of copula theory, known as Sklar's theorem, asserts that any multivariate CDF can be formulated as a combination of its univariate marginal distributions and a copula. Conversely, a copula when applied to univariate margins yields a multivariate CDF. By employing a variant of this theorem adapted for SFs, as introduced by McNeil and Nešlehová (2009), the multivariate SF of $\mathbf{T}$ can be expressed as follows:

$$\bar{F}_{\mathbf{T}}(t_1, \ldots, t_n) = C(\bar{F}_{T_1}(t_1), \ldots, \bar{F}_{T_n}(t_n)), \tag{6}$$

where C is a copula, i.e., a multivariate CDF defined on $[0,1]^n$. This representation is crucial for our analysis because it directly uses SFs, aligning with the transformed Strauss model.

Lemma 1. *The transformed Strauss model, whose survival function is given by Eq.* (2)*, in which the function ψ is involved, admits a Sklar representation (see Eq. 6) through an archimedean copula, expressed as follows:*

$$C(\varphi_1, \ldots, \varphi_n) = \psi\left(\sum_{i=1}^{n} \psi^{-1}(\varphi_i)\right), \tag{7}$$

where ψ is the generator of the copula and ψ^{-1} is its inverse.

Proof. Let $\varphi_i = \bar{F}_{T_i}(t_i)$ for $i = 1, \ldots, n$, where $\bar{F}_{T_i}$ is the marginal survival function of T_i, given by Eq. (4). We thus have:

$$\varphi_i = \psi(\omega_i t_i).$$

Applying the inverse function ψ^{-1} to both sides of this equation gives us:

$$\omega_i t_i = \psi^{-1}(\varphi_i).$$

The multivariate survival function of the transformed Strauss model is given by Eq. (2):

$$\bar{F}_T(t_1, \ldots, t_n) = \psi\left(\sum_{i=1}^{n} \omega_i t_i\right).$$

Substituting $\omega_i t_i$ with $\psi^{-1}(\varphi_i)$, we obtain:

$$\bar{F}_T(t_1, \ldots, t_n) = \psi\left(\sum_{i=1}^{n} \psi^{-1}(\varphi_i)\right).$$

Thus, using the Sklar representation provided by Eq. (6), the associated copula can be expressed as:

$$C(\varphi_1, \ldots, \varphi_n) = \psi\left(\sum_{i=1}^{n} \psi^{-1}(\varphi_i)\right).$$

This proves that the model admits a representation through an archimedean copula whose generator is ψ, thus concluding the proof. □

The expression given by Eq. (7) corresponds to a *strict multivariate archimedean copula*, as defined in the copula literature (see, e.g., Nelsen,

2006), with ψ as the generator.[4] By returning to the utility vector $\mathbf{U}$ of the original Strauss model, the relationship between the survival function $\bar{F}_{\mathbf{T}}$ and the CDF $F_{\mathbf{U}}$ is given by:

$$F_{\mathbf{U}}(x_1, \ldots, x_n) = \bar{F}_{\mathbf{T}}(e^{-x_1/\sigma}, \cdots, e^{-x_n/\sigma}).$$

Similarly, the marginals of $\mathbf{U}$ are related to those of $\mathbf{T}$ by:

$$F_{U_i}(x_i) = \bar{F}_{T_i}(e^{-x_i/\sigma}), \quad i = 1, \ldots, n.$$

Using the representation provided by Eq. (6), we can express $F_{\mathbf{U}}(x_1, \ldots, x_n)$ in terms of the archimedean copula applied to the marginals. This yields the Sklar representation for the CDF of $\mathbf{U}$, according to the following formula:

$$F_{\mathbf{U}}(x_1, \ldots, x_n) = C(F_{U_1}(x_1), \ldots, F_{U_n}(x_n)), \tag{8}$$

where C is exactly the same archimedean copula as identified in Eq. (7).[5]

We will now use this representation of the Strauss model from the copula literature to establish the following result:

Theorem 1. *Consider a model that satisfies the base assumption (Assumption 1). The model is valid if and only if Assumption 2 is also satisfied.*

Proof. We rely on Sklar's representation of the CDF of $\mathbf{U}$, given by Eq. (8). As discussed earlier in this section, the marginals F_{U_i} are valid under assumption 1. However, for the archimedean copula C linking these marginals

[4] In general, as in Nelsen (2006), ψ^{-1} is often referred to as the generator and ψ its inverse. We adopt the same definition as McNeil and Neslehova McNeil and Nešlehová (2009), using ψ as the generator because it slightly simplifies the expressions in our context.

[5] The copula C remaining invariant after a monotone transformation is not specific to this model. In Nelsen (2006), it is shown that when monotone transformations—though not all identical—are applied to the marginals, the copula remains invariant, particularly for the representation of multivariate CDFs with strictly increasing transformations. However, we have not identified a directly applicable result that would allow us to omit these straightforward lines of proof in our specific context.

to be valid as well, it is both necessary and sufficient for the generator function ψ to be completely monotone, according to Assumption 2 along with its corresponding equation, Eq. (5). This result is stated in Nelsen (2006, Theorem 4.6.2, p. 152). For a complete and rigorous proof, the reader is referred to Kimberling (1974), which is referenced in Nelsen's book for a detailed probabilistic interpretation of complete monotonicity. □

In the following section, we will use existing generators from the literature to develop specific models that conform to the Strauss model framework, while ensuring they meet the established assumptions.

4. Examples of Distributions

In this section, we present several examples that illustrate the use of various valid generators ψ within the Strauss model framework. Although our focus is on these specific cases, there is a wide range of other valid generators ψ to construct sarchimedean copulas. For a comprehensive overview of archimedean copulas with a single parameter, readers are referred to (Nelsen, 2006, Table 4.1, pp. 116-118), which lists a broad selection of valid generators. This table serves as a valuable resource for identifying and applying different copula families in various contexts.

A suitable function for ψ within this model is $\psi(y) = \exp(-y)$, which leads to independent double-exponential (Gumbel) distributions. This yields the following multivariate CDF for utilities:

$$F_{\boldsymbol{U}}(x_1, \ldots, x_n) = \exp\left(-\sum_{i=1}^{n} e^{\frac{v_i - x_i}{\sigma}}\right).$$

This formulation of the CDF is consistent with the traditional approach for deriving logit-type choice probabilities.

The marginal distributions of the utilities within this model follow a double-exponential form, each characterized by a common scale parameter.

Using the generator $\psi(y) = \exp(-y)$, whose inverse is $\psi^{-1}(\varphi) = -\ln\varphi$, the resulting copula, derived from substitution in Eq. (7), is:

$$C(\varphi_1, \ldots, \varphi_n) = \prod_{i=1}^{n} \varphi_i,$$

which corresponds to the independent copula. This shows that, in the Strauss model, using $\psi(y) = \exp(-y)$ results in an independent copula where the utilities of different alternatives are independent, each with the same scale parameter.[6]

It is noteworthy to observe, and to draw a conclusion regarding this specific case, that the transformed Strauss model associated with the independent copula possesses the SF:

$$\bar{F}_{\boldsymbol{T}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} e^{-\omega_i t_i}.$$

The model indicates that the resultant distributions are independent exponential distributions, where ω_i serves as the inverse of a scale parameter.

An alternative functional form for ψ that extends the prior independent copula generator is $\psi(y) = \exp(-y^{\beta})$, given that $\beta \in (0, 1]$. Consequently, the CDF is expressed as:

$$F_{\boldsymbol{U}}(x_1, \ldots, x_n) = \exp\left[-\left(\sum_{i=1}^{n} e^{\frac{v_i - x_i}{\sigma}}\right)^{\beta}\right].$$

The standard model, assuming independence is recovered when $\beta = 1$. For $\beta < 1$, the parameter β induces correlations among the alternatives while

[6] It is expected that with independent utilities, the marginal distributions must be double-exponential. Yellott (1977) proved that the multinomial logit model, in the case of independent utilities with at least three alternatives, arises exclusively when the marginal distributions are double-exponential.

maintaining the multinomial logit model of the choice probabilities, as will be elucidated in the subsequent section. Furthermore, it is important to note that this distribution is referenced in (Balakrishnan, 1992, p. 244), where it is used to formulate multivariate logistic distributions.

An additional valid generator ψ within this context is $\psi(y) = (1+y)^{-1}$, which results in a multivariate logistic distribution. Employing this function, the CDF of the utility vector $\mathbf{U}$ is expressed as:

$$F_{\boldsymbol{U}}(x_1, \ldots, x_n) = \frac{1}{1 + \sum_{i=1}^{n} e^{\frac{v_i - x_i}{\sigma}}}.$$

The selection of that ψ induces dependence among the alternatives, in contrast to the independent double-exponential distributions observed with $\psi(y) = \exp(-y)$. In this instance, the utilities cease to be independent, thereby introducing correlations among the alternatives, all while maintaining the multinomial logit model of the choice probabilities.

It should be noted that in the well-known handbook on the logistic distribution, namely Balakrishnan (1992), a generalization is provided with distinct scale parameters σ_i the last equation displayed above regarding the multivariate logistic distribution. Nonetheless, as previously noted, this generalization would undermine the logit and exploded logit forms of the choice and ranking probabilities, respectively. Specifically, the integrands within the integrals lack explicit primitives unless the scale parameters are uniform across all alternatives. Consequently, it is imperative to retain a singular σ to uphold these two fundamental properties.

5. Exploded Logit

In this section, we examine the probability of the event in which the alternatives are ordered in strictly decreasing utility, denoted as $(U_1 > U_2 > \cdots > U_n)$, within the context of the Strauss model. We wish to calculate the

probability of this event for our purpose, namely:

$$\mathbb{r}(\omega_1, \cdots, \omega_n) \equiv P(U_1 > U_2 > \ldots > U_n) = P(T_1 < \ldots < T_n).$$

It is crucial to emphasize that this specific ordering does not compromise the generality of our analysis. Despite the existence of $n!$ potential ranking probabilities, we focus exclusively on a single order to examine its analytical form. The indexing of alternatives is, in fact, arbitrary and the derived results are applicable to any permutation of indices. This formulation enables us to determine all other ranking probabilities by permuting the indices. In doing so, we simplify our notation and facilitate a clearer and more concise presentation of the results.

The calculation of ranking probabilities is more demanding than that of best-choice probabilities, which we will briefly revisit here by illustrating how to determine, without loss of generality, the probability that the labeled alternative 1 has a utility greater than those of the other alternatives. To achieve this, our objective is to calculate the best-choice probabilities defined as follows:

$$\mathbb{b}(\omega_1, \cdots, \omega_n) \equiv P(U_1 > U_i,\ i = 2, \ldots, n) = P(T_1 < T_i,\ i = 2, \ldots, n).$$

This notation indicates the probability that the alternative represented by the first argument in the vector is preferred over the other alternatives labeled in the vector. We use the same precaution as before regarding the notation $\mathbb{r}(\omega_1, \ldots, \omega_n)$ and the fact that the number of arguments indicates the number of alternatives in consideration.

The calculation of ranking probabilities is more demanding than that of best-choice probabilities, which we will briefly revisit here by illustrating how to determine, without loss of generality, the probability that the labeled 1 has a utility greater than those of the other alternatives.

We thus need to calculate the following integral:

$$\mathbb{b}(\omega_1, \cdots, \omega_n) = \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_1}^\infty f_{\boldsymbol{T}}(t_1, \ldots, t_n)\, dt_n \ldots dt_1, \tag{9}$$

where $f_{\boldsymbol{T}}$ is the joint density of the utility vector $\boldsymbol{T}$, given by:

$$\begin{aligned} f_{\boldsymbol{T}}(t_1, \ldots, t_n) &\equiv (-1)^n \frac{\partial^n \bar{F}_{\boldsymbol{T}}(t_1, \ldots, t_n)}{\partial t_1 \ldots \partial t_n} \\ &= (-1)^n \left(\prod_{i=1}^n \omega_i \right) \psi^{(n)} \left(\sum_{i=1}^n \omega_i t_i \right). \end{aligned} \tag{10}$$

Using the fact that the density function $f_{\boldsymbol{T}}$, as shown in Eq. (10), is obtained through successive derivatives of the survival function, and applying the fundamental theorem of calculus $n-1$ times, we can simplify the expression given by Eq. (9) to obtain a simpler integral that depends only on t_1. This approach is valid because all integral bounds are fixed at t_1, which does not depend on $t_2, \ldots, t_n$. We then need to calculate:

$$\mathbb{b}(\omega_1, \cdots, \omega_n) = -\int_0^\infty \frac{\partial \bar{F}_{\boldsymbol{T}}(t_1, \ldots, t_1)}{\partial t_1}\, dt_1,$$

which yields:

$$\mathbb{b}(\omega_1, \cdots, \omega_n) = -\omega_1 \int_0^\infty \psi' \left(\sum_{k=1}^n \omega_k t_1 \right) dt_1 = \frac{\omega_1 [\psi(0) - \psi(\infty)]}{\sum_{k=1}^n \omega_k}.$$

Using Assumption 1, we obtain:

$$\mathbb{b}(\omega_1, \cdots, \omega_n) = \frac{\omega_1}{\sum_{i=1}^n \omega_i} = \frac{e^{v_1/\sigma}}{\sum_{i=1}^n e^{v_i/\sigma}}. \tag{11}$$

This proves that the Strauss model is compatible with the multinomial logit model for best-choice probabilities. This relationship expands the possibilities regarding distributions, beyond those of utilities limited to independent

double-exponential distributions, thus leading to these forms of best-choice probabilities.

Ranking probabilities are more complex and can be computed using the following multiple integrals:

$$\mathbb{r}(\omega_1, \ldots, \omega_n) = \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty f_{\boldsymbol{T}}(t_1, \ldots, t_n)\, dt_n \ldots dt_1. \tag{12}$$

Unlike the integrals that arise in the calculation of best-choice probabilities, here the bounds of the $n-1$ inner integrals involve all the arguments $t_1, \ldots, t_{n-1}$ rather than just t_1. This makes it impossible to use the fundamental theorem of calculus recursively to reduce the calculation to a simple integral after $n-1$ applications of the theorem. However, as we will see later, in the case of the Strauss model, the calculation will not be so insurmountable and will lead to an explicit form.

By inserting the expression for the density given in Eq. (10) into the multiple integral that appears in Eq. (12), we obtain the following form:

$$\mathbb{r}(\omega_1, \ldots, \omega_n) = \prod_{i=1}^n \omega_i \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty (-1)^n \psi^{(n)} \left(\sum_{i=1}^n \omega_i t_i \right) dt_n \ldots dt_1. \tag{13}$$

Strauss (1979) proved that the distribution of utilities he proposed, which we have termed the Strauss model, conforms to the exploded logit framework, offering a more general proof compared to that of Beggs et al. (1981), which is constrained to independent double-exponentials. These two papers utilized backward integration in their respective proofs; in contrast, our approach introduces an alternative proof method.

Instead of performing backward integration for an alternative ranked before the others, starting from the last-ranked alternative and working backward, we use an inductive approach based on the number of alternatives. This methodology facilitates a direct transition from a set comprising n alternatives to a set that includes $n+1$ alternatives, eliminating the need

to handle integrals sequentially. Additionally, we have established robust conditions for the Strauss model using strict archimedean copulas, thereby strengthening our contribution with solid conditions for the Strauss model and an alternative proof.

Theorem 2 (Strauss, 1979)**.** *Assume that the utility vector* $\mathbf{U}$ *follows the CDF of the Strauss model as specified in Eq.* (1)*, and that the function* ψ *in this CDF satisfies Assumptions 1 and 2. Under these conditions, the ranking probability* $\mathbb{r}$ *is given by the exploded logit model and is described by the following formula:*

$$\mathbb{r}(\omega_1, \ldots, \omega_n) = \prod_{i=1}^{n-1} \frac{\omega_i}{\sum_{j=i}^{n} \omega_j} = \prod_{i=1}^{n-1} \frac{e^{v_i/\sigma}}{\sum_{j=n}^{n} e^{v_j/\sigma}}. \tag{14}$$

Proof. The case $n = 2$ corresponds to the calculation of the probability $P(T_1 < T_2)$ and is aligned with a binary choice probability. We can then apply Eq. (11) in this particular case to obtain:

$$\mathbb{r}(\omega_1, \omega_2) = \mathbb{b}(\omega_1, \omega_2) = \frac{\omega_1}{\omega_1 + \omega_2}.$$

By induction hypothesis, suppose the equality holds for all integers up to some $n \geq 2$. We need to show that this remains true for $n + 1$. Using Eq. (13), we need to compute:

$$\begin{aligned}\mathbb{r}(\omega_1, \ldots, \omega_{n+1}) = &\prod_{i=1}^{n+1} \omega_i \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_n}^\infty \\ &(-1)^{n+1} \psi^{(n+1)} \left(\sum_{i=1}^{n+1} \omega_i t_i \right) dt_{n+1} \ldots dt_1.\end{aligned}$$

We start by integrating with respect to t_{n+1}, which gives:

$$\begin{aligned}\mathbb{r}(\omega_1,\ldots,\omega_{n+1}) =& \prod_{i=1}^{n}\omega_i \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty \\ & (-1)^n \psi^{(n)}\left(\sum_{i=1}^{n-1}\omega_i t_i + (\omega_n+\omega_{n+1})t_n\right) dt_n \ldots dt_1.\end{aligned}$$

It is useful to express this last equation in the following form:

$$\begin{aligned}\mathbb{r}(\omega_1,\ldots,\omega_{n+1}) =& \frac{\omega_n}{\omega_n+\omega_{n+1}} \\ & (\omega_n+\omega_{n+1})\prod_{i=1}^{n-1}\omega_i \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{n-1}}^\infty \\ & (-1)^n \psi^{(n)}\left(\sum_{i=1}^{n}\omega_i t_i + (\omega_n+\omega_{n+1})t_n\right) dt_n \ldots dt_1.\end{aligned}$$

Using Eq. (13), this expression can be simplified to:

$$\mathbb{r}(\omega_1,\ldots,\omega_{n+1}) = \frac{\omega_n}{\omega_n+\omega_{n+1}}\mathbb{r}(\omega_1,\cdots,\omega_{n-1},\omega_n+\omega_{n+1}).$$

We then apply the induction hypothesis to obtain, after simplification:

$$\mathbb{r}(\omega_1,\ldots,\omega_{n+1}) = \prod_{i=1}^{n}\frac{\omega_i}{\sum_{j=i}^{n+1}\omega_j},$$

which completes the proof. □

To make a final remark before moving to the conclusion of our article, we would like to highlight the term "exploded," which illustrates how the probabilistic ranking model obtained from the Strauss model arises from the decomposition of ranking probabilities into a product of best-choice probabilities. At each step, we evaluate the probability that the alternative in the first position is the best among all n alternatives. We then eliminate this

alternative and assess the probability that the next alternative, in the second position, is the best among the remaining alternatives labeled $2, \ldots, n$. This process continues until only one alternative remains, numbered n, while maintaining the order from 1 to n.

Using our notations, we can express Eq. (14) from Theorem 2 as follows:

$$\mathbb{r}(\omega_1, \cdots, \omega_n) = \prod_{i=1}^{n-1} \mathbb{b}(\omega_i, \cdots, \omega_n).$$

This relationship is obtained through the equations of the multinomial logit model, which the Strauss model adheres to, as presented in Eq. (11), defining the best-choice probabilities.

6. Conclusions and research perspectives

In this paper, we revisit the Strauss model by delving into the necessary and sufficient conditions for its validity using modern copula theory. By specifying the minimal assumptions on the support of marginal distributions and proving that the Strauss model can be represented by strict archimedean copulas, we have clarified and strengthened the underlying structure of this random utility model.

A major contribution of our work is the rigorous formulation of Assumptions 1 and 2, which ensure the validity of the multivariate CDF used in the Strauss model. Assumption 2 requires that the generator function ψ of the archimedean copula be completely monotone, which means that its successive derivatives strictly alternate in sign. This property, known as *complete monotonicity*, is essential to ensure that all marginal and joint densities derived from ψ meet the positivity conditions necessary for the validity of the model. By incorporating the theory of strict archimedean copulas, we have shown that the Strauss model naturally fits within this specific class of copulas, thereby facilitating the analysis of dependencies between the utilities of different alternatives.

A particularly interesting aspect of our analysis is the clarification that the Strauss model generates the exploded logit model commonly used in econometric applications. Strauss proved that his model, more generally than that often subsequently attributed to Beggs et al. (1981), leads to the exploded logit model. Contrary to what is frequently found in the literature, where this model is associated with independent utilities, our study highlights that the utilities in the Strauss model can be dependent and cannot be considered independent without altering the model's fundamental properties. In reality, the independent case of the Strauss model corresponds only to double-exponentials, which represents a small part of the broader class of distributions considered. As soon as a different copula is used, generating a copula generator other than the independent one, other specifications of dependence emerge. This distinction reaffirms Strauss's pioneering contribution and cautions against the systematic association of the exploded logit model with independent utilities.

Furthermore, by examining different copula specifications, we proved that only independent double-exponentials lead to the exploded logit model. The use of different copulas generates other dependence specifications, limiting the ability to generalize the exploded logit model while maintaining explicit analytical properties. It is crucial to understand these dependence structures between the utilities of alternatives, as they directly influence the ability to propose models alternative to the exploded logit. If the goal is to construct models that are not exploded logit and offer only advantages, it is necessary to avoid dependence structures and marginal types characteristic of the Strauss model. An additional contribution of our work is the proposal of a new proof of the exploded logit model by induction, thus reinforcing the results established by Strauss.

For future research, it would be relevant to explore generalizations that allow for distinct scale parameters, which would extend the Strauss model. Although maintaining a single scale parameter σ is essential to preserve the

exploded logit form, introducing different scale parameters eliminates the possibility of explicit analytical forms. However, this extension remains interesting for numerical methods, though it goes beyond the scope of this paper. Additionally, incorporating strict archimedean copulas with multiple parameters or other copula families could allow for modeling more nuanced dependencies between the utilities of alternatives, thus offering greater flexibility to adapt to various empirical contexts. This extension could enhance the model's ability to capture complex relationships observed in real data.

In conclusion, our study provides a theoretical clarification of the Strauss model and proves its potential for extension through the integration of strict archimedean copulas. It also underscores the importance of recognizing Strauss's pioneering contributions to the development of random utility models. These advances offer new opportunities to improve discrete choice modeling and better capture dependencies between the utilities of alternatives. Ultimately, we hope this research will assist in addressing the open question discussed in this article, namely: Are there random utility models for which the best choice probabilities satisfy the multinomial logit model, but the ranking probabilities do not follow the exploded logit model?"

Acknowledgements

I would like to express my sincere thanks to Hans Colonius for his support and valuable advice, and to Paolo Delle Site, Stéphane Hess, Thomas Hancock, Jean Lainé, and Michel Regenwetter for their encouragement and the enriching discussions I have had with them, either in person or through correspondence, on various occasions regarding discrete choice models.

References

Balakrishnan, N., 1992. Handbook of the Logistic Distribution. Statistics, Textbooks and Monographs ; v. 123, Dekker, New York.

Beggs, S., Cardell, S., Hausman, J., 1981. Assessing the potential demand for electric cars. Journal of Econometrics 17, 1–19.

Joe, H., 1997. Multivariate Models and Multivariate Dependence Concepts. 1st ed., Chapman and Hall/CRC, Boca Raton, FL.

Kimberling, C.H., 1974. A probabilistic interpretation of complete monotonicity. Aequationes mathematicae 10, 152–164.

Luce, R.D., 1959. Individual Choice Behavior: A Theoretical analysis. Wiley, New York, NY, USA.

McNeil, A.J., Nešlehová, J., 2009. Multivariate Archimedean copulas, d-monotone functions and 1-norm symmetric distributions. The Annals of Statistics 37, 3059 – 3097.

Nelsen, R.B., 2006. An Introduction to Copulas. Second ed., Springer, New York, NY, USA.

Schwiebert, J., 2016. Multinomial choice models based on archimedean copulas. AStA Advances in Statistical Analysis 100, 333–354.

Strauss, D., 1979. Some results on random utility models. Journal of Mathematical Psychology 20, 35–52.

Yellott, J.I., 1977. The relationship between luce's choice axiom, thurstone's theory of comparative judgment, and the double exponential distribution. Journal of Mathematical Psychology 15, 109–144.