Journal of Choice Modelling 21 (2016) 25–29

Contents lists available at ScienceDirect

Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Presenting the Independence of Irrelevant Alternatives property in a first course on logit modeling

Patricia L. Mokhtarian

School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA, 30332 United States

ARTICLE INFO

Keywords:
Endogeneity
Independence of Irrelevant Alternatives (IIA)
Logit models
Omitted variables bias

ABSTRACT

The brief treatment of IIA offered in most introductory courses and textbooks leaves a number of important points unsaid, and may paint an incomplete and misleading picture of its nature. This note supplements such treatments with additional detail. Specifically, we first present three causes or types of IIA violations, and describe some common ways in which they occur. We then present some typical tests/remedies for IIA violations, indicating which types of violation they are best suited to identify, and how they "work" in terms of remedying a violation. We close with some additional pertinent observations.

1. Introduction

One of the first things students learn about the multinomial logit (MNL) model is that it possesses the Independence of Irrelevant Alternatives (IIA) property, meaning that the relative probability of choosing one alternative over another is independent of what other alternatives are in or out of the choice set. A useful corollary is that cross-(disaggregate point) elasticities are uniform, meaning that the percentage change in the probability of choosing alternative i, given a percentage change in attribute m of alternative j, is constant for all $i \neq j$ (Ben-Akiva and Lerman, 1985).[1]

The sometimes unrealistic nature of this property is often illustrated by the "Red Bus/Blue Bus (RB/BB) paradox", and instructors generally indicate that the reason for the IIA violation is that the assumption of independent error terms is clearly untenable in this instance[2]. Nested logit may be portrayed as one test for IIA violations, and a solution if IIA *is* violated by the MNL model, precisely because it allows the error terms of "similar" alternatives to be correlated.

The brief treatment of IIA offered in most introductory courses and textbooks leaves a number of important points unsaid, and may paint an incomplete and misleading picture of its nature. The purpose of this short note is to supplement such treatments with helpful additional detail, drawn from a variety of published and unpublished sources. Specifically, we first present three causes or types of IIA violations, and describe some common ways in which they occur. We then present some typical tests/remedies for IIA violations, indicating which types of violation they are best suited to identify, and how they "work" in terms of remedying a violation. We close with some additional pertinent observations.

2. Three sources of IIA violation

To establish (the entirely standard) notation, let

E-mail address: patmokh@gatech.edu.

[1] As a reminder, these features of IIA only apply at the disaggregate level; even for MNL, IIA does not hold in the aggregate.

[2] For models other than MNL, neither "IIA holds" nor "independent and identically-distributed error terms" necessarily implies the other (Bjorn and Vuong, 1985).

http://dx.doi.org/10.1016/j.jocm.2016.08.002
Received 30 September 2015; Received in revised form 5 July 2016; Accepted 31 August 2016
Available online 22 September 2016
1755-5345/

$$U_{jn} = V_{jn} + \varepsilon_{jn},$$

where U_{jn} is the utility of alternative *j* to decision-maker *n*;

V_{jn} is the observed portion of utility, generally characterized as a linear-in-parameters function $\boldsymbol{\beta}'\mathbf{x}_{jn}$, where $\boldsymbol{\beta}$ is a vector of parameters and $\mathbf{x}_{jn}$ is a conformable vector of attributes of the alternative and the individual; and

ε_{jn} is the unobserved portion of utility.

Assuming that V_{jn} is independent of ε_{jn}, and that the ε_{jn}s are independent and identically extreme-value distributed (i.i.d. EV, and setting the scale parameter of the distribution to one for ease of exposition, since the scale parameter cannot be separately identified and the magnitude of $\boldsymbol{\beta}$ adjusts depending on the value assumed for that parameter), leads to the MNL probability that person *n* chooses alternative *i* (Ben-Akiva and Lerman, 1985):

$$P_n(i) = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}}.$$

Satisfaction of the IIA property requires that V_{in} be a function only of attributes of alternative *i*, not of any other alternatives. Otherwise, if it were a function of, say, attributes of alternative *j*, then $P_n(i)/P_n(k)$ would differ depending on whether alternative *j* were in the choice set or not.

In part because of the attention given to the RB/BB paradox in illustrating the nature of an IIA violation, many people are left with the impression that IIA violations are a matter of how "similar" or "correlated" two alternatives are. This is faulty for two reasons. First, alternatives per se are not correlated; only *measurements* of a specific characteristic or variable associated with an alternative can be correlated (just as it is meaningless to talk about an "average person", without specifying the characteristic(s) over which the average is being taken: age? height? income?).[3] Second, as we will see more clearly below, whether or not IIA is violated "depends on the specification of the model for a particular choice situation, not on the choice situation itself" (McFadden et al., 1977, p. 45). Thus, deducing a violation of IIA must be a matter of empirical testing on a particular model specification, not a presumption based on the nature of the available alternatives. The literature (McFadden et al., 1977; Ben-Akiva and Lerman, 1985) contains multiple examples in which, *for the same set of alternatives*, one model specification violates IIA while another one does not.

Even if students understand this point, however, they are likely to draw the also-erroneous conclusion that *the only* reason for IIA to be violated is because unobserved characteristics of alternatives are shared, meaning that the error terms ε_{in} and ε_{jn} of the respective utility functions are correlated. For example, Ben-Akiva and Lerman (1985, p. 109) note that "the core of the problem is the assumption that the disturbances are mutually independent", although they may mean "the problem" to refer specifically to the RB/BB paradox, not to the IIA problem in general.

However, although it is seldom noted explicitly in introductory treatments, it is important to realize that IIA will also be violated if ε_{in} is not independent of V_{in} or of V_{jn} (e.g., McFadden et al., 1977, p. 41: "Generally, the violations may be traced to the MNL assumption that the unobserved-utility component is independent across alternatives *and independent of the observed attributes*", emphasis added). Thus, we can define three common types of violations of IIA[4]:

Type 1: $\boldsymbol{\varepsilon_{in}}$ is correlated with $\boldsymbol{\varepsilon_{jn}}$. This is by far the most commonly-discussed type. It is exemplified in the extreme by the RB/BB example, in which many unobserved characteristics relevant to choice are shared by red buses and blue buses.

Type 2: $\boldsymbol{\varepsilon_{in}}$ is correlated with $\mathbf{V_{in}}$. The derivation of MNL probabilities (Ben-Akiva and Lerman, 1985) assumes that V_{in} is uncorrelated with ε_{in}. If actually V_{in} and ε_{in} are correlated, so that we could write $V_{in}= \alpha_0+\alpha_1\ \varepsilon_{in}+\xi_{in}$ for some $\alpha_1\neq 0$ and random perturbation ξ_{in} (Tardiff, 1979), then V is clearly *not* independent of ε, and the distribution of V + ε is more complicated than assumed.

Type 3: $\boldsymbol{\varepsilon_{jn}}$ is correlated with $\mathbf{V_{in}}$. Again, in the derivation of MNL probabilities we assume that $U_{in}=V_{in} + \varepsilon_{in} \sim EV(V_{in}, \mu)$ and $U_{jn}=V_{jn}+\varepsilon_{jn}\sim EV(V_{jn}, \mu)$ (where μ is the scale parameter conventionally fixed at 1 for the identifiability of $\boldsymbol{\beta}$) are *independent* as well as identically distributed (the latter meaning both EV-distributed, with the same scale parameter μ). Only then will the difference of EV-distributed random variables have the logistic distribution, as is necessary for MNL to result. If $corr(V_{in}, \varepsilon_{jn})\neq 0$, then most likely $corr(U_{in}, U_{jn})\neq 0$, and the independence assumption is violated. In any case, if $corr(V_{in}, \varepsilon_{jn})\neq 0$, so that we can write $V_{in}=\gamma_0+\gamma_1\ \varepsilon_{jn}+\tau_{in}$ for some $\gamma_1\neq 0$ and random perturbation τ_{in}, then just as for Type 2, the distribution of $U_{in}-U_{jn}=V_{in}-V_{jn} + \varepsilon_{in} - \varepsilon_{jn}$ is more complicated than we have assumed. We will also violate the condition that

$$P_n(i)/P_n(k) = \exp(V_{in}-V_{kn}) = \exp(\gamma_0 + \gamma_1\varepsilon_{jn} + \tau_{in}-V_{kn})$$

should depend only on characteristics of alternatives *i* and *k* – it will also depend on (unobserved) characteristics of alternative *j* (Tardiff, 1979).

In linear regression analysis, Type 2 and 3 violations are examples of a familiar specification error, namely the exclusion of relevant variables, in the sinister case in which the excluded variables (ε_{in} and/or ε_{jn}) are correlated with included ones ($\mathbf{x}_{in}$). A number of early references (e.g. Tardiff, 1979; Horowitz, 1981; Lee, 1982) address this type of specification error in the context of discrete choice modeling, noting that in that context, parameter estimates will be biased or inconsistent or both. For example, if the

[3] For the same reason, alternatives cannot be "independent" of each other per se, and thus the very phrase "Independence of Irrelevant Alternatives" contributes to the confusion.

[4] IIA will also be violated if $Var(\varepsilon_{in})$ is heteroscedastic (violating the "identically-distributed" assumption); this type is not treated here.

unobserved degree of luxury possessed by a transportation mode is correlated with the observed cost of taking that mode, clearly the estimated coefficient of cost in a discrete choice model will be partly accounting for the effect of luxury, and would therefore reflect a biased estimate of the effect of cost per se.

More broadly, the Type 2 and 3 violations constitute an *endogeneity bias* (Wooldridge, 2010, pp. 54–55), which can arise due to *omitted variables* (as described in the previous paragraph), *measurement error* (where the difference between the true and the observed measurement can be viewed as an omitted variable), or *simultaneity* (in which, if an observed explanatory variable x_m is an effect as well as a cause of U, x_m will likely be a function of – thence correlated with – the ε of U).

There is a sizable literature on endogeneity in the context of discrete choice models, generally completely unlinked to a discussion of IIA. Although it is beyond the scope of this brief note to discuss in detail, a number of advanced remedies forendogeneity have been developed (e.g. Train, 2009, Chapter 13; Louviere et al., 2005; de Grange et al., 2015), including the Berry/Levinsohn/Pakes (BLP) approach, instrumental variable or control functions, the multiple imputations approach to addressing measurement error (Brownstone et al., 2001), hybrid choice models (Vij and Walker, 2016), and multiplicative error models (Gandhi et al., 2010).

Returning to the three typical ways in which IIA can be violated, it is important to note that we are not ordinarily concerned about V_{in} being correlated with V_{jn}, which probably happens more often than not. There are at least three common ways in which this could happen:

a. **Characteristics are conceptually related between alternatives**. For example, because travel times (TTs) are related to distance, and the distance is (roughly) the same across alternatives in a given mode choice context, TT_{bus} and TT_{car} are probably correlated across *n*: when travel time by bus is high for a person *n*, travel time by car will also tend to be.
b. **There may not be a conceptual relationship, but an empirical one**. For example, the comfort of transit may theoretically have nothing to do with the comfort of cars. But if, across *n*, cars tend to be perceived as comfortable and buses tend to be perceived as uncomfortable, measured values of $comfort_{car}$ and $comfort_{bus}$ could be (negatively) correlated.
c. **Decision-maker-specific variables associated with the utilities of multiple alternatives will be (perfectly) correlated.**

If none of the types of IIA violation occur, correlations of V_{in} and V_{jn} can safely be neglected. If any of the violations do occur, however, such correlations become more problematic. For example, if a Type 2 violation occurs, in which ε_{in} is correlated with V_{in}, then if V_{in} is also correlated with V_{jn}, V_{jn} may be correlated with ε_{in} (a Type 3 violation). If, in addition, ε_{jn} is correlated with V_{jn}, (which is likely, for the same reason that ε_{in} would be correlated with V_{in}) then not only may ε_{jn} be correlated with V_{in} (by virtue of the correlation between V_{in} and V_{jn}) – another Type 3 violation – but ε_{jn} may also be correlated with ε_{in} (by virtue of the "chain" $Corr(\varepsilon_{jn}, V_{jn})$, $Corr(V_{jn}, V_{in})$, $Corr(V_{in}, \varepsilon_{in})$) – a Type 1 violation. Of course, theoretically, two variables may each be highly correlated with a third but not at all correlated with each other (Langford et al., 2001), and empirically, the longer the "chain" of correlations linking two variables, in general the weaker the latter correlation becomes. Nevertheless, it seems prudent to suspect that if one source of violation is present, others are likely to be as well. Conversely, "breaking the chain" by removing one source of violation may (but is not guaranteed to) resolve other types of violation as well.

Now, consider that the ways listed above, in which the utilities for alternatives *i* and *j* could be correlated, can also apply to the *unobserved εs* as well as to the *observed Vs*. When that is the case, correlations of ε_{in} and ε_{jn}, a Type 1 IIA violation, will result. With respect to the first way, for example, if person *n* has a general attitude toward transit, then her specific (unobserved) attitudes regarding the convenience, comfort, status, safety, etc. of bus will probably be correlated with her counterpart attitudes toward rail – a classic reason for a Type 1 violation. With respect to the third way, if an unobserved individual-specific characteristic differentially affects the utility of multiple alternatives, its presence in the error terms of two or more alternatives will also generate a Type 1 violation.

With respect to Type 2 violations, one common basis for them is the correlation of an observed variable with a conceptually different unobserved variable, as in the above example of cost and luxury. Another common basis, however, is a specification that constrains coefficients for a given attribute to be equal across alternatives, when in fact the coefficient differs across alternatives. To see this, suppose that $\beta_k x_{ink}$ and $\beta_k x_{jnk}$ should really be $\beta^*_{ik} x_{ink}=(\beta_k+ b_{ik}) x_{ink}$ and $\beta^*_{jk} x_{jnk}=(\beta_k+ b_{jk}) x_{jnk}$, respectively, where the alternative-specific coefficient β^*_{ik} can be decomposed into a mean (across alternatives) component β_k and an alternative-specific deviation from the mean b_{ik}, and similarly for β^*_{jk}. Then in the constrained model a Type 2 violation will result, because V_{in} and ε_{in} will be correlated through the common presence of x_{ink} in $\beta_k x_{ink}$ for V_{in} and $b_{ik} x_{ink}$ for ε_{in} (and similarly for alternative j). Furthermore, if x_{ink} and x_{jnk} are correlated, through any of the mechanisms (a) through (c) presented above, then a Type 1 violation also results in the constrained model, since ε_{in} and ε_{jn} are correlated because of the presence of $b_{ik} x_{ink}$ in ε_{in} and $b_{jk} x_{jnk}$ in ε_{jn}.

3. Common tests and remedies for IIA violations

A full discussion of the numerous tests and remedies for IIA violations is beyond the scope of this short note. We will select some basic tests/remedies that could easily be presented in an introductory course, and focus on their relationship to the three types of violation introduced in the previous section. Table 1 shows eight of the most common tests/remedies, distinguished by their rationale (Fry and Harris, 1996) – either to compare MNL with an alternative, more general model specification that does not embody the IIA assumption (where, if a statistical test rejects the null hypothesis that the two models are equivalent, the remedy is

Table 1
Common tests and remedies for IIA violations, by type of violation addressed.

Test/remedy	How it works	Type of violation addressed		
		1. ε_{in} correlated with ε_{jn}	**2. ε_{in} correlated with V_{in}**	**3. ε_{jn} correlated with V_{in}**
Comparison to a more general model, namely …				
Universal logit (Hausman and McFadden, 1984)	Adds the relevant variables from V_{in} into V_{jn}, thereby removing them from ε_{jn}			√
Add variables important to choice	Moves them from ε_{in} to V_{in}; can include individual-specific traits common to multiple εs	(√) (e.g. if variable is individual-specific)	√ if added vars. corr. with vars. already included in V_{in}	(√) (e.g. if variable is indiv.-specific, and corr. w/V)
Add variables that measure traits common to some alternatives but not others	Moves those characteristics from ε_{in} and ε_{jn} to V_{in} and V_{jn}, respectively	√		
Change variables from generic to alternative- specific	When $\beta^*_{ik}=\beta+b_{ik}$ and $\beta^*_{jk}=\beta+b_{jk}$, moves "$b_{ik}x_{ink}$" from ε_{in} to V_{in}, and similarly for j	(√) (if x_{ink} and x_{jnk} are correlated)	√ thru x_{ink}, which appears in both V_{in} (βx_{ink}) and ε_{in} ($b_{ik}x_{ink}$)	(√) (if x_{ink} and x_{jnk} are correlated)
Dogit (Gaudry and Dagenais, 1979)	Vs from all alternatives allowed to affect relative probability of choosing i over j			√
Nested logit	Allows εs to be correlated across alternatives	√		
Non-i.i.d. probit	Allows εs to be correlated across alternatives	√		
Comparison of coefficients across different choice sets				
Hausman-McFadden (Hausman and McFadden, 1984)	If there is correlation of ε_{jn} with ε_{in} or V_{in}, estimated coefficients will differ depending on whether *j* is in or out of the choice set	√		√

√: This test/remedy effective with the indicated violation.
Note: In view of the discussion in the text, a √ for any one type of IIA violation could also imply √s for the other two, *if* the relevant variables are correlated accordingly. Rather than fill the table with √s, I have only highlighted some cases where the possible correlation of variables seems particularly germane.

typically to adopt the more general model); or to compare the coefficients of a model estimated on the full choice set with those of a counterpart estimated on a reduced choice set (where, if a statistical test rejects equivalence, the remedy is to search for a new specification). The table could readily be extended to accommodate additional tests/remedies found in the literature.

A key point is that a particular test/remedy may be suited to detect/resolve one form of IIA violation but not others. For example, if the problem is that two alternatives share relevant unobserved characteristics (Type 1), then bringing variables from V_{jn} into V_{in} (as in universal logit or dogit) may do nothing to resolve the violation. Conversely, if the problem is that unobserved relevant variables are correlated with included variables (Types 2 and 3), then turning to a nested logit or non-i.i.d. probit model to allow error terms to be correlated will do nothing to resolve the issue (Horowitz, 1981).

Thus, to be more confident that IIA is not violated by a given MNL model, it is important to test for all three types of violations. Even so, of course such a process is not foolproof. For one thing, a Type II statistical error could occur: failing to reject the null hypothesis of IIA holding when in fact it is false. But even if the hypothesis test gives the “correct” answer, none of the remedies are guaranteed to work. For example, adding variables important to choice will not help a Type 2 violation if the added variables (moved from ε to V) were not correlated with V to begin with – the variables remaining in ε will still be correlated with V. Thus, as with all statistical testing, there are elements of chance and judgment.

4. Conclusion

The purpose of this note is to clarify that, contrary to the stereotype perpetuated by the Red Bus, Blue Bus paradox, IIA violations can be of three common types. They arise not only when unobserved portions of utility are correlated across alternatives, but also when the unobserved portion of utility for a given alternative is correlated with the observed portion of utility for the same or for a different alternative (i.e., when there is an endogeneity bias, whether due to omitted variables, measurement error, or simultaneity). The note presents common tests/remedies for IIA violations in the light of explaining which type(s) of violation they are best-suited to identify, and how they detect and/or remedy a violation. I stress that a specific test for IIA may detect one or two of the three types of violation but not the other(s), suggesting that multiple tests for IIA violations should be conducted so as to ensure that all three types of violation are checked.

Acknowledgment

The constructive comments of two anonymous reviewers were very helpful in improving this paper.

References

Ben-Akiva, Moshe, Lerman, Steven R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.
Bjorn, Paul, Vuong, Quang H., 1985. A note on the independence of irrelevant alternatives in probabilistic choice models. Economics Letters 18, 305–307.
Brownstone, D., Golob, T., Kazimi, C., 2001. Modeling non-ignorable attrition and measurement error in panel surveys. In: Groves, R., Dillman, D., Eltinge, J., Little, R. (Eds.), Ch. 25 in Survey Nonresponse. Wiley, New York, 373–388.
Fry, Tim R.L., Harris, Mark N., 1996. A Monte Carlo study of tests for the Independence of Irrelevant Alternatives property. Transp. Res. B 30 (**1**), 19–30.
Gandhi Amit, Kim Kyoo il and Petrin Amil, 2010. The Interaction of Observed and Unobserved Factors in Discrete Choice Demand Models. Unpublished Working Paper. Available at ⟨http://apps.olin.wustl.edu/MEGConference/Files/pdf/2010/72.pdf⟩ (accessed 04.07.16).
Gaudry, Marc J.I., Dagenais, M.G., 1979. The dogit model. Transp. Res. 13B, 105–112.
de Grange, Louis, Gonzalez, Felipe, Vargas, Ignacio, Troncoso, Rodrigo, 2015. A logit model with endogenous explanatory variables and network externalities. Netw. Spat. Econ. 15, 89–116.
Hausman, Jerry, McFadden, Daniel, 1984. Specification tests for the multinomial logit model. Econometrica 52 (**5**), 1219–1240.
Horowitz, Joel, 1981. Identification and diagnosis of specification errors in the multinomial logit model. Transp. Res. 15B, 345–360.
Langford, Eric, Schwertman, Neil, Owens, Margaret, 2001. Is the property of being positively correlated transitive? Am. Stat. 55 (4), 322–325.
Lee, Lung-Fei, 1982. Specification error in multinomial logit models: analysis of the omitted variable bias. J. Econ. 20, 197–209.
Louviere, Jordan, Train, Kenneth, Ben-Akiva, Moshe, Bhat, Chandra, Brownstone, David, Cameron, Trudy Ann, Carson, Richard T., DeShazo, J.R., Fiebig, Denzil, Greene, William, Hensher, David, Waldman, Donald, 2005. Recent progress on endogeneity in choice modeling. Mark. Lett. 16 (3/4), 255–265.
McFadden, D., Train, K., Tye, W., 1977. An application of diagnostic tests for the independence of irrelevant alternatives property of the multinomial logit model. Transp. Res. Rec. 637, 39–46.
Tardiff, Timothy J., 1979. Specification analysis for quantal choice models. Transp. Sci. 13 (3), 179–190.
Train, Kenneth, 2009. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK, Available at: ⟨http://elsa.berkeley.edu/~train/distant.html⟩.
Vij, Akshay, Walker, Joan L., 2016. How, when and why integrated choice and latent variable models are latently useful. Transp. Res. B 90, 192–217.
Wooldridge, Jeffrey M., 2010. Econometric Analysis of Cross Section and Panel Data 2nd ed.. MIT Press, Cambridge, MA.