HARVARD | BUSINESS | SCHOOL

Taste Heterogeneity, IIA, and the Similarity Critique

Thomas J. Steenburgh
Andrew Ainslie

Working Paper

09-049

Taste Heterogeneity, IIA, and the Similarity Critique

Thomas J. Steenburgh and Andrew Ainslie

Abstract

The purpose of this paper is to show that allowing for taste heterogeneity does not address the similarity critique of discrete-choice models. Although IIA may technically be broken in aggregate, the mixed logit model allows neither a given individual nor the population as a whole to behave with perfect substitution when facing perfect substitutes. Thus, the mixed logit model implies that individuals behave inconsistently across choice sets.

Estimating the mixed logit on data in which individuals do behave consistently can result in biased parameter estimates, with the individuals' tastes for desirable attributes being systemically undervalued.

Key words: Heterogeneity, Mixed Logit, Independence from Irrelevant Alternatives, IIA, Similarity Critique, Ecological Fallacy

[1] The authors would like to thank David E. Bell and Sunil Gupta for comments and suggestions that have greatly improved the paper. The authors alone, of course, are responsible for remaining errors.

1 Introduction

It is widely believed that choice models that allow for taste heterogeneity overcome the issues created by the Independence from Irrelevant Alternatives (IIA) property. It has been argued that allowing for differences in tastes *across* individuals creates correlations in the random component of utility that "break" the undesirable IIA property and allow any substitution pattern to be found. Thus, the problems with IIA are thought to be solved by models such as the mixed logit [cf. Allenby and Lenk, 1994, Brownstone and Train, 1998, Revelt and Train, 1998, McFadden and Train, 2000, Glasgow, 2001, Chintagunta et al., 2003, Erdem et al., 2008].

This belief raises some conceptually troubling questions: if a model implies that individuals make inconsistent choices across different presentations of a choice set when their tastes are assumed to be the same, why would the same model predict that individuals behave consistently when their tastes are assumed to be different? On the other hand, if the substitution patterns are fixed only in the population, should we not be concerned that this is another example of the ecological fallacy [Robinson, 1950] or Simpson's paradox [Simpson, 1951]?[2] Even if the model does allow more flexible substitution to be found, how do we explain why allowing for differences in tastes solves the problem?

It seems to have been lost in the current literature that the critiques of Debreu [1960], Savage [Luce and Suppes, 1965] and McFadden [1974], among others, suggest more than that IIA is an undesirable property. These critiques also suggest the type of behavior that is needed to ensure that individuals behave consistently across choice sets, specifically that they behave with *perfect substitution*[3] when facing perfect substitutes. It has not been shown that models such as the mixed logit allow this to occur.

The purpose of this paper is to show that allowing for taste heterogeneity does not address the similarity critique of discrete-choice models, either at the individual or at the aggregate level of a model. IIA may not technically hold for the population if tastes vary across individuals, but the mixed logit model allows neither a given individual nor the population as a whole to behave with perfect substitution when facing perfect substitutes. Thus, the mixed logit implies that individuals behave inconsistently across different presentations of the same choice set. Furthermore, estimating the mixed logit on data in

[2] The ecological fallacy and Simpson's paradox both describe situations in which an incorrect conclusion about individual behavior is reached by observing population behavior.

[3] We use the term *perfect substitution* to describe the rational choice behavior suggested in each of the similarity critiques. This will be discussed in greater detail in Section 2.

which individuals do behave consistently can result in biased parameter estimates, with the individuals' tastes for desirable attributes being systemically undervalued.

2 The Similarity Critique

2.1 Individual Choice Behavior

Let us recall the similarity critique by considering an example proposed by Steenburgh [2008]. Suppose an individual faces a choice between two laptop computers:

	Weight	Processor Speed
Laptop A	3 lbs.	2.0 GHz
Laptop B	6 lbs.	3.0 GHz

Laptop A is the lighter alternative, but it runs at a slower speed; Laptop B is faster, but heavier. The laptops are identical in all other ways. The individual is indifferent between the two laptop computers and would choose either with probability $1/2$.

Now suppose a third alternative, Laptop B′, is added to the choice set. Laptop B′ is identical to Laptop B; it weighs 6 lbs and runs at 3.0 GHz. Thus, the individual would be equally likely to choose either Laptop B or Laptop B′ if she were asked to choose between just the two of them. What would happen if the individual were asked to choose among all three alternatives? If the individual were to behave consistently across the two choice sets, Laptop A would be chosen with probability $1/2$, Laptop B with probability $1/4$, and Laptop B′ with probability $1/4$. This is the desirable, rational behavior that we refer to as perfect substitution.

But a model with IIA precludes this from happening because it requires the ratio of any two choice probabilities to remain the same no matter what other alternatives are included in the choice set. According to a model with IIA, the individual would choose Laptop A with probability $1/3$, Laptop B with probability $1/3$, and Laptop B′ with probability $1/3$. The individual does not behave consistently across the two choice sets because the probability that a faster laptop is chosen rises from $1/2$ in the original set to $2/3$ in the set with a perfect substitute and the probability that the lighter laptop is chosen falls from $1/2$ to $1/3$. Thus, the individual behaves irrationally.[4]

While it is commonly understood that IIA is an undesirable property because it results in irrational choice behavior, it seems to be overlooked that simply breaking IIA is not sufficient to ensure that individuals behave consistently. To clarify this idea, suppose a third model implies that Laptop A would be chosen with probability $1/5$, Laptop B with probability $2/5$, and Laptop B′ with probability $2/5$ when Laptop B′ is added to the choice set. This new model clearly breaks IIA. At the same time, it is even more objectionable than the one with IIA because the probability that a faster laptop is chosen rises from $1/2$ in the original set to $4/5$ when the perfect substitute is added and the probability that the lighter laptop is chosen falls from $1/2$ to $1/5$. These represent even greater changes in probability.

A good choice model should allow rational choice behavior to occur, not simply break IIA. This means that individuals need to be able to behave with perfect substitution when facing perfect substitutes.

2.2 The Similarity Critique and Population Choice Behavior

It has been shown that IIA does not hold for the population if individuals have heterogeneous tastes; thus, it has been supposed that IIA is not a concern as long as a researcher's interest centers around only the population's choice behavior. By extending the previous example, we will show that while taste heterogeneity may break IIA at the population level of a model, it does not truly address the similarity critique.[5]

Consider the following example. Suppose there are two types of individuals in the population, Salespeople and Scientists, and the proportion of individuals in the population of each type is $1/2$. Both Salespeople and Scientists prefer laptop computers that weigh less and that run faster. Nevertheless, Salespeople value lighter weights more than Scientists do, and Scientists value faster processor speeds more than Salespeople do. Suppose that when presented with a choice

---

[4] Rational choice behavior, a cornerstone of micro-economic theory, requires an individual to be able to rank all of the alternatives in order of preference and an individual's preferences to be consistent. Thus, an individual cannot be indifferent between A and B in one choice set, yet prefer B to A in another. Obviously, by slightly modifying the example it can be shown that the logit model implies strict preference reversals [Debreu, 1960].

[5] Even if allowing for taste variation did address the similarity critique for the population, it would be somewhat troubling to rely on a paradox to do so [Steenburgh, 2008], especially in micro-targeting applications in which individual choice behavior must be understood.

between Laptops A and B, a Salesperson chooses Laptop A with probability $^2/_3$ whereas a Scientist chooses Laptop A with probability $^1/_3$.

The market share of each laptop computer is a property of the population. It can be thought of as the probability that an individual chosen at random from the population chooses a given laptop computer. Thus, the market share of Laptop A is $\Pr\{A\} = \Pr\{A|Saleperson\}\Pr\{Saleperson\} + \Pr\{A|Scientist\}\Pr\{Scientist\}$. Taken together, these assumptions imply the following choice probabilities:

	Salespeople	Scientists	Population
Laptop A	$^2/_3$	$^1/_3$	$^1/_2$
Laptop B	$^1/_3$	$^2/_3$	$^1/_2$

Again suppose a third alternative identical to Laptop B, weighing 3.0 lbs and running at 2.0 GHz, is added to the choice set. If the individuals were to behave with perfect substitution, the following choice probabilities and market shares would occur:

	Salespeople	Scientists	Population
Laptop A	$^2/_3$	$^1/_3$	$^1/_2$
Laptop B	$^1/_6$	$^1/_3$	$^1/_4$
Laptop B′	$^1/_6$	$^1/_3$	$^1/_4$

Yet, if both Salespeople and Scientists were to behave with IIA, the following choice probabilities and market shares would arise:

	Salespeople	Scientists	Population
Laptop A	$^1/_2$	$^1/_5$	$^{14}/_{40}$
Laptop B	$^1/_4$	$^2/_5$	$^{13}/_{40}$
Laptop B′	$^1/_4$	$^2/_5$	$^{13}/_{40}$

Strictly speaking, it should be clear that the market shares do not possess IIA. If they did, the new market shares would be exactly $^1/_3$ for Laptop A, $^1/_3$ for Laptop B, and $^1/_3$ for Laptop B′, as they were for the individual in the previous example. Nevertheless, it should also be clear that allowing for differences in tastes does not address the similarity critique in aggregate. The population does not behave consistently across the two choice sets because the market share of the faster laptop increases from from $^1/_2$ in the original set to $^{14}/_{20}$ (or 65%) in the set with a perfect substitute and the share of the lighter laptop falls from $^1/_2$ to $^{14}/_{40}$ (or 35%).[6] Thus, the similarity critique has not truly been addressed.

The following proposition asserts that this is true in general.

Proposition 1 *If individuals behave with IIA, overall demand for a given alternative* must *increase if a perfect substitute for it is added to the choice set. This occurs for every individual in the population, and therefore occurs for the population as a whole, regardless of whether individuals' tastes are different.*

PROOF. It suffices to consider a choice between two alternatives. Suppose individual n chooses one alternative with probability P_n and a composite of other alternatives with probability $(1 - P_n)$. Make no assumptions about the individuals' tastes, so the P_n can vary across individuals.

Suppose a perfect substitute for the first alternative is added to the choice set. Since individuals behave with IIA, the new choice probabilities satisfy $k_n P_n + k_n P_n + k_n (1 - P_n) = 1$, which implies $k_n = {}^1/_{(1+P_n)}$.

An individual is more likely to pick the alternative with a perfect substitute in the expanded choice set if $2k_n P_n > P_n$. This occurs if $P_n < 1$, which is obviously true.

Since every individual is more likely to choose the alternative with a perfect substitute, the market share of that alternative must increase too.

$\sum\limits_{\forall n} 2k_n P_n > \sum\limits_{\forall n} P_n$. QED[7]

3 Population Correlations and Individual Choice Behavior

In this section, we will derive the mixed logit from the standard random-coefficients perspective, will discuss the assumptions that are made at the individual and population levels of the model, and will draw a parallel with the earlier work of Robinson [1950].

[6] In fact, the population's behavior is much closer to IIA than it is to perfect substitution.

[7] The authors thank David E. Bell for suggesting a greatly refined proof.

3.1 *Individual Choice Behavior*

Let us begin with the standard theory of how individuals choose among discrete alternatives. Suppose an individual faces a choice among J alternatives. The utility obtained by individual n from alternative j on a given occasion is

$$\begin{gathered} u_{nj} = x_j^{'}\beta_n + \varepsilon_{nj} \\ \varepsilon_{nj} \overset{iid}{\sim} EV\left(I\right)\ \ for\ j = 1, \ldots, J \end{gathered} \tag{1}$$

The individual's utility is separated into two components. The first component, $x_j\beta_n$, is referred to as the observed utility. It is a function of the observed attributes of each alternative, x_{jt}, and the individual's tastes for those attributes, β_n. The second component, ε_{nj}, is referred to as the unobserved utility. It is a random variable that accounts for factors other than the alternative's attributes that affect the individual's utility.

The unobserved utilities are commonly assumed to be independent and identically distributed across alternatives according to a type-I extreme value distribution. Thus, they are uncorrelated at the individual level of the model.

$$Cov\left(\varepsilon_{nj}, \varepsilon_{nk}\right) = 0 \quad \forall j \neq k \tag{2}$$

The link between the individual's utility and observed choice behavior is established as follows. Suppose the individual chooses the alternative that provides the greatest utility. The decision rule governing his or her behavior is to choose alternative j if and only if $u_j > u_k\ \forall k \neq j$. Given the assumptions of the model, the probability that individual n chooses alternative j conditional on his or her tastes is

$$\begin{aligned} P_{nj} &= \Pr\left\{\varepsilon_{nk} - \varepsilon_{nj} < x_j\beta_n - x_k\beta_n\ \forall k \neq j \left|\beta_n\right.\right\} \\ &= \frac{\exp\left(x_j\beta_n\right)}{\sum\limits_{\forall k}\exp\left(x_k\beta_n\right)} \end{aligned} \tag{3}$$

Thus, the basic model implies that individuals behave with IIA.

3.2 *Population Choice Behavior*

The theory of individual choice is extended to the population by adding an additional level to the model. In the mixed logit, the utilities of all individuals in the population are jointly modeled as

$$\begin{aligned} u_{nj} &= x_j \beta_n + \varepsilon_{nj} \\ \varepsilon_{nj} &\overset{iid}{\sim} EV(I) \quad \text{for } n = 1, \ldots N \text{ and } j = 1, \ldots, J \\ \beta_n &\overset{iid}{\sim} MVN(\theta, \Sigma) \quad \text{for } n = 1, \ldots N \end{aligned} \tag{4}$$

The unobserved utilities for any individual are assumed to be independent across alternatives. Thus, the probability that individual n chooses alternative j remains as expressed in equation 3 and individuals behave with IIA [cf. Erdem et al., 2008]. [8]

The mixed logit can be transformed by expressing the individuals' tastes in terms of differences from the mean tastes in the population. Defining $\delta_n \equiv \beta_n - \theta$, the model can be written as

$$\begin{aligned} u_{nj} &= x_j \theta + x_j \delta_n + \varepsilon_{nj} \\ \varepsilon_{nj} &\overset{iid}{\sim} EV(I) \;\; for\, n = 1, \ldots N \; and\, j = 1, \ldots, J \\ \delta_n &\overset{iid}{\sim} MVN(0, \Sigma) \;\; for\, n = 1, \ldots N \end{aligned} \tag{5}$$

By considering the variation in tastes *across* individuals in addition to the variation at the individual level, we can define a random component of utility as $\eta_{nj} \equiv x_j \delta_n + \varepsilon_{nj}$. A correlation between alternatives exists in this random component of utility, as the covariance between η_{nj} and η_{nk} is

$$\begin{aligned} \sigma_{\cdot jk} &= Cov(\eta_{nj}, \eta_{nk}) \\ &= x_j' \Sigma x_k \end{aligned} \tag{6}$$

It should clear, however, that this correlation is a property of the population as opposed to a property of an individual because it is determined by integrating over the tastes of all individuals in the population. In contrast, the random component of utility at the individual level, ε_{nj} and ε_{nk}, remains uncorrelated between alternatives as stated in equation 2.

[8] Some have gone so far as to imply that the mixed logit model breaks IIA at the individual level. The choice probabilities that they suggest [cf. Train, 2003, p. 142], however, are integrated over the tastes of all individuals in the population and therefore are not individual choice probabilities.

It has been conjectured that the population correlations overcome the problems with IIA, but neither an individual (equation 3) nor the population as a whole (Proposition 1) behaves rationally in a mixed logit model. This leads to a number of questions: Even if the population correlations did completely address the issues with IIA, how do we give economic meaning to this solution in anything other than a technical sense? Does it matter that the covariation between alternatives depends only on Σ, which is a property of the population as opposed to the individual? Why does the presence of individuals with different tastes affect the substitution behavior of any given individual? Does it matter that the mixed logit model collapses into the multinomial logit if tastes do not vary across individuals?

The previous work of Robinson [1950] should make us even more cautious about inferring that population (ecological) correlations can fix a problem with individual choice behavior. Robinson shows that finding correlations in a population does not necessarily imply that correlations exist among individuals. A parallel can be drawn with the mixed logit in that the correlations between alternatives exist only in aggregate. Robinson also suggests that the behavior of a population should not be applied to the behavior of individuals[9], and we may wonder whether the mixed logit can fit the data well if individuals do behave rationally.

4 Monte Carlo Choice Experiments

A test for whether the mixed logit model allows for perfect substitution can be constructed by estimating it across multiple Monte Carlo experiments in which the alternatives are presented both with and without perfect substitutes. If the choice behavior is captured by the model, then the parameter estimates will be consistently recovered across the experiments. If not, then the parameter estimates will systematically vary. We will show that estimating the mixed logit model on data that contain a perfect substitute results in biased parameter estimates.

4.1 Population of Individuals with Homogeneous Tastes

To build intuition for how the taste estimates are affected, first consider a population of hypothetical individuals with homogeneous tastes. Suppose that these individuals are asked to choose among different configurations of laptop computers in three separate Monte Carlo experiments. The individuals face

[9] Robinson's critique has come to be known as the ecological fallacy.

the same series of choices in each experiment, but the presentation of the alternatives varies across experiments. The laptops have varying weights and processor speeds, but are identical in all other ways.

In the first Monte Carlo experiment, individuals are presented with choices between two alternatives: one that dominates on weight and another that dominates on processor speed. The similarity critique does not apply to the choice behavior in this experiment because any two alternatives are equally similar to each other; thus, choices can be simulated by assuming that the individuals behave according to a multinomial logit model.

In the next two Monte Carlo experiments, individuals face the same series of choices, but three alternatives are presented in each choice set instead of two. In one experiment, a perfect substitute for the alternative that dominates on weight is added to the set; in the other, a perfect substitute for the alternative that dominates on processor speed is added. Individuals are assumed to behave rationally. Thus, the choice probability for an alternative with a perfect substitute is assumed to be exactly $1/2$ of what it would have been if the perfect substitute were not available. For example, suppose an individual chooses a laptop that weighs 3 lbs and runs at 2.0 GHz with probability $1/2$ in a given choice set. If a perfect substitute for that alternative is added to the set, the choice probabilities for these two alternatives each would be $1/4$ and the remaining choice probabilities would all stay the same.

Suppose the population consists of 100 individuals whose choices, when only two alternatives are presented, can be described by a multinomial logit model with tastes for weight and speed of {-0.3470, 1.041} respectively. These tastes imply that a 1 lb reduction in a laptop's weight would produce the same change in utility as a $1/3$ GHz increase in its processor speed for each individual in the population. For simplicity, nine different sets of alternatives were used to simulate data. These sets along with the corresponding choice probabilities for each alternative in each of the Monte Carlo experiments are presented in Table 1.

In order to eliminate concerns about sample size, ninety choices were simulated for each individual (ten repetitions of the nine choice sets), which resulted in a total of 9,000 simulated choices. Given the design of the experiment, roughly half of the observed choices should have been for the alternative that dominated on weight and half for the alternative that dominated on processor speed when the data are pooled across all choice sets. This was observed in the data, as the alternative that dominated on weight was respectively chosen 49.4%, 49.6% and 50.1% of the time across the three Monte Carlo experiments. A summary of the observed choices is provided in Table 2.

The researcher's estimation problem is to draw inference about the individuals' tastes from their observed choices. To accomplish this, we estimated a multinomial logit model[10] using the observed choice data in each of the three Monte Carlo experiments. A summary of the results is presented in Table 3.

Although the data were generated by assuming that individuals behave consistently across choice sets, the parameter estimates vary widely across the three Monte Carlo experiments. The results are reasonable when individuals are presented with only two alternatives. The multinomial logit model estimates that a 1 lb reduction in a laptop's weight would produce the same change in utility as a 0.341 GHz increase in its processor speed, which is close to the true value of 0.333 GHz.

Nevertheless, the model produces undesirable results when individuals face perfect substitutes: it underestimates the individuals' tastes for the attribute on which the perfect substitutes dominate. When estimated on the data set that contains perfect substitutes for the lightweight alternative, the multinomial logit model estimates that a 1 lb reduction in the laptop's weight produces the same change in utility as a 0.123 GHz increase in its processor speed. The model predicts that individuals value lighter weights much less than they actually do. Conversely, when estimated on a data set that contains perfect substitutes for the fast-processor alternative, the multinomial logit model estimates that a 1 lb reduction in the laptop's weight produces the same change in utility as a 1.14 GHz increase in processor speed. The model predicts that individuals value faster processor speeds much less than they actually do.

What causes these two estimates to differ by nearly an order of magnitude? The multinomial logit model undervalues the individuals' tastes for the dominant attribute of the perfect substitutes because utility is constructed as a scalar index that ignores the similarity of the alternatives in attribute space. A perfect substitute must draw half of an individual's choices away from its twin alternative if an individual is to behave consistently across choice sets. The model cannot capture this behavior because the similarity of the alternatives is not accounted for either in the utility function or elsewhere in the model. Thus, when a perfect substitute for the alternative that dominates on weight is available, the model attributes the lower frequency of its choice to individuals valuing lighter weights less than they actually do rather than to a greater number of alternatives being available. Likewise, when a perfect substitute for the alternative that dominates on processor speed is available, the model attributes its lower frequency of choice to individuals valuing faster processors less than they actually do rather than to a greater number of alternatives being available.

[10] All models were estimated using bayesm, an open software package available for the R programming language. See Rossi et al. [2005] for details.

4.2 Population of Individuals with Heterogeneous Tastes

Now consider a population of individuals with heterogeneous tastes that is subjected to the same experiment. Suppose this population consists of two types of people, Salespeople and Scientists. The Salespeople's tastes differ from the Scientists', but within each subpopulation the tastes of individual Salespeople and individual Scientists are the same. The Salespeople's tastes for weight and speed respectively are {-0.4618, 0.6927}, which implies that a 1 lb. reduction in weight produces the same change in utility as a 0.667 GHz increase in processor speed. The Scientists' tastes for weight and speed respectively are {-0.2322, 1.3932}, which implies that a 1 lb. reduction in weight produces the same change in utility as a 0.167 GHz increase in clock speed. The choice sets along with the corresponding choice probabilities for both types of individuals are presented in Table 4.

Suppose the population consists of 100 Salespeople and 100 Scientists. Every individual in the population makes 90 choices, which results in a total of 18,000 observations. Given the design of the experiment, Salespeople should have chosen the laptop that dominated on weight roughly 61% of the time whereas Scientists should have chosen it roughly 39% of the time. This was observed in the data as Salespeople chose the laptop that dominated on weight 61.4%, 62.0% and 61.8% of the time whereas Scientists chose this alternative 37.8%, 39.4% and 39.5% of the time across the three Monte Carlo experiments. A summary of the observed choices is provided in Table 5.

To infer the individuals' tastes from their observed choices, we estimated the following random coefficients logit model:

$$\begin{gathered} u_{njt} = x_{jt}\beta_n + \varepsilon_{njt} \\ \varepsilon_{njt} \overset{iid}{\sim} EV\left(I\right) \quad \text{for } n = 1, \ldots N,\ j = 1, \ldots, J \text{ and } t = 1, \ldots, T \\ \beta_n \overset{iid}{\sim} N\left(\Delta' z_n, \Sigma\right) \quad \text{for } n = 1, \ldots N \end{gathered} \tag{7}$$

where the demographic variable z_n indicates whether an individual is a Salesperson or a Scientist.

Numerical summaries of the results are presented in Tables 6 and 7. Table 6 contains the parameter estimates and Table 7 contains the implied tradeoffs between weight and processor speed for the average Salesperson and the average Scientist. Graphical summaries of how tastes are distributed across individuals within each subpopulation are provided in Figure 1; these box plots are constructed by using the 100 posterior means of β_n within each subpopulation as data. The labels EXP I, EXP II, and EXP III refer to the Monte Carlo experiments in which individuals respectively faced a choice set containing only two alternatives, perfect substitutes for the lightweight alternative, and perfect substitutes for the fast-processor alternative.

Although the individuals are assumed to behave consistently no matter how the alternatives are presented in the choice set, the estimates of their tastes again vary widely across the three Monte Carlo experiments. The results are reasonable when individuals face only two alternatives. The mixed logit model estimates that a 1 lb reduction in weight would produce the same change in utility as a 0.656 GHz increase in processor speed for the average Salesperson and as a 0.150 GHz increase for the average Scientists, which are close to the true values. Also, the distributions of individual tastes for both Salespeople and Scientists surround the true parameter values that were assumed in the experiment, as can be seen in the box plots for EXP I.

Nevertheless, the mixed logit model produces biased parameter estimates when individuals face perfect substitutes: it underestimates both the Salespeople's and the Scientists' tastes for the attribute on which the perfect substitutes dominate. When estimated on the data set that contains a perfect substitute for the lightweight alternative (EXP II), the mixed logit model estimates that a 1 lb reduction in the laptop's weight produces the same change in utility as a 0.255 GHz increase in processor speed for the average Salesperson and as a 0.0420 GHz increase for the average Scientist. The model predicts that both Salespeople and Scientists value lighter weights much less than they actually do. Conversely, when estimated on a data set with a perfect substitutes for the fast-processor alternative (EXP III), the model estimates that a 1 lb reduction in the laptop's weight produces the same change in utility as a 1.64 GHz increase in processor speed for the average Salesperson and as a 0.502 GHz increase for the average Scientist. The model predicts that both Salespeople and Scientists value faster processor speeds much less than they actually do.

It may be thought that the mixed logit model is able to recover the distribution of tastes in the population even if it fails to recover the tastes of any individual correctly, but this is not the case either. The mixed logit model undervalues the tastes of both Salespeople and Scientists for the attribute on which a perfect substitute dominates, which shifts the distributions of both subpopulations in the same direction. As can been seen in the box plots, the distributions of both the Salespeople's and the Scientists' tastes for weight shifts toward zero (meaning that weight is less valued) and for speed shifts away from zero (meaning that speed is more highly valued) in EXP II. Conversely, the opposite occurs in EXP III.

These Monte Carlo experiments show that allowing for differences in tastes across individuals does not address the similarity critique. The presence of a perfect substitute biases the parameter estimates in the same direction for all individuals under study, with the mixed logit model undervaluing both the Salespeople's and the Scientists' tastes for the dominant attribute of the perfect substitute.

4.2.1 Policy Analysis

Choice models are commonly used to predict the market shares that would arise following a new product introduction. Consumers' tastes are elicited through choice-based conjoint experiments, various new configurations of alternatives are proposed, and market shares are simulated using the choice model. Obviously, this procedure is not useful if the parameter estimates are biased, but we might wonder if concerns still exist if only two alternatives are presented in the conjoint experiment. Unfortunately, this is not the case because the model does not allow the desired substitution patterns to be predicted when a new alternative is introduced.

Consider the market shares predicted by the mixed logit model in EXP I for the laptops in the opening example. (These shares are reported in Table 8.) The model makes accurate predictions in Choice Set I, which contains only two alternatives, just as the data used to estimate the model did. The model, however, mispredicts the market shares when a third alternative is introduced. When a perfect substitute for the lighter alternative is introduced (Choice Set II), the market share of the lighter laptop rises from 49.6% to 64.0% in the population. Conversely, when a perfect substitute for the faster alternative is introduced (Choice Set III), the market share of the faster laptop rises from 50.4% to 64.8% in the population. Similar substitution patterns are found in the market segments too.

The implication for marketers is clear: the mixed logit model would over-predict the market share that arises for a me-too product if it is introduced into the choice set, would under-predict the share stolen from the replicated incumbent, and would over-predict the share stolen from the distinct incumbent.

5 Conclusion

It is undoubtedly important to account for differences in tastes if we want to study people's choices in almost any real world setting, and the modeling innovations of the past several decades will continue to prove useful in this regard. Nevertheless, this study definitively shows that allowing for differences in tastes is not adequate to address the similarity critique.

A choice model with IIA implies that individuals behave inconsistently and therefore irrationally across choice sets. In prediction, it implies that overall demand for a given alternative must rise if a perfect substitute for it is added to the choice set, which paints an overly optimistic picture for me-too product introductions. In estimation, parameters will be biased if rational behavior is observed, with the taste of every individual being systematically undervalued.

The heart of the matter is that IIA is an assumption about how an individual behaves, and concerns about it cannot be addressed by allowing for differences across individuals.

References

Greg M. Allenby and Peter J. Lenk. Modeling household purchase behavior with logistic normal regression. *Journal of the American Statistical Association*, 89(428):1218–1231, 1994. ISSN 01621459. URL http://www.jstor.org/stable/2290986.

David Brownstone and Kenneth Train. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89(1-2):109–129, November 1998.

P.K. Chintagunta, J.P. Dubé, and V. Singh. Balancing Profitability and Customer Welfare in a Supermarket Chain. *Quantitative Marketing and Economics*, 1(1):111–147, 2003.

Gerard Debreu. Review of individual choice behavior: A theoretical analysis by r. duncan luce. *The American Economic Review*, 50(1):186–188, Mar 1960.

T. Erdem, M.P. Keane, and B. Sun. The impact of advertising on consumer price sensitivity in experience goods markets. *Quantitative Marketing and Economics*, 6(2):139–176, 2008.

Garrett Glasgow. Mixed logit models for multiparty elections. *Political Analysis*, 9(2):116–136, 2001. URL http://pan.oxfordjournals.org/cgi/content/abstract/9/2/116.

R. Duncan Luce and Patrick Suppes. Preference, utility, and subjective probability. In R. Duncan Luce, Robert R. Bush, and Eugene Galanter, editors, *Handbook of Mathematical Psychology*, volume III, pages 249–410. John Wiley & Sons, New York, 1965.

D. McFadden and K. Train. Mixed mnl models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka, editor, *Frontiers in Econometrics*. Academic Press, 1974.

David Revelt and Kenneth Train. Mixed logit with repeated choices: Households' choices of appliance efficiency level. *The Review of Economics and Statistics*, 80(4):647–657, 1998. ISSN 00346535. URL http://www.jstor.org/stable/2646846.

William S. Robinson. Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3):351–357, Jun 1950.

Peter E. Rossi, Greg M. Allenby, and Robert McCulloch. *Bayesian Statistics and Marketing*. John Wiley & Sons, 2005.

E. H. Simpson. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(2):238–241, 1951. ISSN 00359246. URL http://www.jstor.org/stable/2984065.

Thomas J. Steenburgh. The invariant proportion of substitution property (ips) of discrete-choice models. *Marketing Science*, 27(2):300–307, March-April 2008.

Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, U.K., 2003.

Table 1
Choice Probabilities for the Population of Individuals with Homogeneous Tastes

Choice Set	Alternatives	EXP I	EXP II	EXP III
I	4.5 lbs, 2.0 GHz	0.5	0.25, 0.25	0.5
	6.0 lbs, 2.5 GHz	0.5	0.5	0.25, 0.25
II	3.0 lbs., 2.0 GHz	0.627	0.314, 0.314	0.627
	6.0 lbs., 2.5 GHz	0.372	0.372	0.186, 0.186
III	4.5 lbs, 2.0 GHz	0.372	0.186, 0.186	0.372
	6.0 lbs, 3.0 GHz	0.627	0.627	0.314, 0.314
IV	3.0 lbs, 2.0 GHz	0.5	0.25, 0.25	0.5
	6.0 lbs, 3.0 GHz	0.5	0.5	0.25, 0.25
V	4.5 lbs, 2.5 GHz	0.5	0.25, 0.25	0.5
	6.0 lbs, 3.0 GHz	0.5	0.5	0.25, 0.25
VI	3.0 lbs, 2.5 GHz	0.627	0.314, 0.314	0.627
	6.0 lbs, 3.0 GHz	0.372	0.372	0.186, 0.186
VII	3.0 lbs, 2.0 GHz	0.5	0.25, 0.25	0.5
	4.5 lbs, 2.5 GHz	0.5	0.5	0.25, 0.25
VIII	3.0 lbs, 2.0 GHz	0.372	0.186, 0.186	0.372
	4.5 lbs, 3.0 GHz	0.627	0.627	0.314, 0.314
IX	3.0 lbs, 2.5 GHz	0.5	0.25, 0.25	0.5
	4.5 lbs, 3.0 GHz	0.5	0.5	0.25, 0.25

Table 2
Observed Choices for the Population of Individuals with Homogeneous Tastes

Choice Set	Alternatives	EXP I	EXP II	EXP III
I	4.5 lbs., 2.0 GHz	491	245, 250	484
	6.0 lbs, 2.5 GHz	509	505	231, 285
II	3.0 lbs., 2.0 GHz	604	314, 291	639
	6.0 lbs., 2.5 GHz	396	395	178, 183
III	4.5 lbs, 2.0 GHz	377	173, 178	403
	6.0 lbs, 3.0 GHz	623	649	288, 309
IV	3.0 lbs, 2.0 GHz	506	249, 260	497
	6.0 lbs, 3.0 GHz	494	491	262, 241
V	4.5 lbs, 2.5 GHz	516	245, 251	518
	6.0 lbs, 3.0 GHz	484	504	234, 248
VI	3.0 lbs, 2.5 GHz	648	289, 341	613
	6.0 lbs, 3.0 GHz	352	370	191, 196
VII	4.5 lbs, 2.5 GHz	483	484	251, 235
	3.0 lbs, 2.0 GHz	517	230, 286	514
VIII	3.0 lbs, 2.0 GHz	365	168, 192	396
	4.5 lbs, 3.0 GHz	635	640	289, 315
IX	3.0 lbs, 2.5 GHz	526	253, 272	473
	4.5 lbs, 3.0 GHz	474	475	262, 265
Totals	Lighter Weight	4450	2166, 2321	4537
	Faster Processor	4550	4513	2186, 2277

Table 3
Parameter Estimates for the Population with Homogeneous Tastes

	Truth	EXP I	EXP II	EXP III
Weight	-0.347	-0.351	-0.194	-0.476
		(0.022)	(0.022)	(0.021)
Speed	1.041	1.027	1.578	0.418
		(0.066)	(0.065)	(0.064)
1 lb weight reduction =	0.333 GHz	0.342 GHz	0.123 GHz	1.14 GHz

Table 4. Choice Probabilities for the Population with Heterogeneous Tastes

		EXP I		EXP II		EXP III	
Choice Set	Alternatives	Salespeople	Scientists	Salespeople	Scientists	Salespeople	Scientists
I	4.5 lbs., 2.0 GHz	0.586	0.414	0.293, 0.293	0.207, 0.207	0.586	0.414
	6.0 lbs, 2.5 GHz	0.414	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293
II	3.0 lbs., 2.0 GHz	0.739	0.500	0.369, 0.369	0.250, 0.250	0.739	0.500
	6.0 lbs., 2.5 GHz	0.261	0.500	0.261	0.500	0.131, 0.131	0.250, 0.250
III	4.5 lbs, 2.0 GHz	0.500	0.260	0.250, 0.250	0.130, 0.130	0.500	0.260
	6.0 lbs, 3.0 GHz	0.500	0.740	0.500	0.740	0.250, 0.250	0.370, 0.370
IV	3.0 lbs, 2.0 GHz	0.667	0.333	0.333, 0.333	0.166, 0.166	0.667	0.333
	6.0 lbs, 3.0 GHz	0.333	0.667	0.333	0.667	0.167, 0.167	0.333, 0.333
V	4.5 lbs, 2.5 GHz	0.586	0.414	0.293, 0.293	0.207, 0.207	0.586	0.414
	6.0 lbs, 3.0 GHz	0.414	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293
VI	3.0 lbs, 2.5 GHz	0.739	0.500	0.369, 0.369	0.250, 0.250	0.739	0.500
	6.0 lbs, 3.0 GHz	0.261	0.500	0.261	0.500	0.131, 0.131	0.250, 0.250
VII	3.0 lbs, 2.0 GHz	0.586	0.414	0.293, 0.293	0.207, 0.207	0.586	0.414
	4.5 lbs, 2.5 GHz	0.739	0.586	0.739	0.586	0.207, 0.207	0.293, 0.293
VIII	3.0 lbs, 2.0 GHz	0.500	0.260	0.250, 0.250	0.130, 0.130	0.500	0.260
	4.5 lbs, 3.0 GHz	0.500	0.740	0.500	0.740	0.250, 0.250	0.370, 0.370
IX	3.0 lbs, 2.5 GHz	0.586	0.414	0.293, 0.293	0.207, 0.207	0.586	0.414
	4.5 lbs, 3.0 GHz	0.414	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293

Table 5. Observed Choices for the Population with Heterogeneous Tastes

		EXP I		EXP II		EXP III	
Choice Set	Alternatives	Salespeople	Scientists	Salespeople	Scientists	Salespeople	Scientists
I	4.5 lbs., 2.0 GHz	592	419	297, 329	203, 206	594	431
	6.0 lbs, 2.5 GHz	408	581	374	591	203, 203	282, 287
II	3.0 lbs., 2.0 GHz	751	507	366, 364	269, 259	769	519
	6.0 lbs., 2.5 GHz	249	493	270	472	123, 108	242, 239
III	4.5 lbs, 2.0 GHz	493	250	236, 255	123, 136	490	262
	6.0 lbs, 3.0 GHz	507	750	509	741	243, 267	347, 391
IV	3.0 lbs, 2.0 GHz	684	314	339, 340	181, 154	662	347
	6.0 lbs, 3.0 GHz	316	686	321	665	165, 173	324, 329
V	4.5 lbs, 2.5 GHz	577	420	303, 287	219, 201	599	411
	6.0 lbs, 3.0 GHz	423	580	410	580	193, 208	294, 295
VI	3.0 lbs, 2.5 GHz	746	458	358, 384	247, 256	736	488
	6.0 lbs, 3.0 GHz	254	542	258	497	139, 125	260, 252
VII	3.0 lbs, 2.0 GHz	581	372	275, 313	196, 211	618	406
	4.5 lbs, 2.5 GHz	419	628	412	593	193, 189	309, 285
VIII	3.0 lbs, 2.0 GHz	498	251	264, 274	119, 135	503	260
	4.5 lbs, 3.0 GHz	502	749	462	746	242, 255	360, 380
IX	3.0 lbs, 2.5 GHz	606	412	302, 291	223, 211	587	430
	4.5 lbs, 3.0 GHz	394	588	407	566	200, 213	279, 291
Totals	Lighter Weight	5528	3403	2740, 2837	1780, 1769	5558	3554
	Faster Processor	3472	5597	3423	5451	1701, 1741	2697, 2749

Table 6. Parameter Estimates for the Population with Heterogeneous Tastes

		Truth	EXP I	EXP II	EXP III
Salespeople	Weight	-0.462	-0.509	-0.322	-0.701
			(0.033)	(0.032)	(0.036)
	Speed	0.693	0.777	1.266	0.428
			(0.085)	(0.079)	(0.087)
Scientists	Weight	-0.232	-0.217	-0.076	-0.390
			(0.035)	(0.032)	(0.034)
	Speed	1.393	1.443	1.817	0.776
			(0.086)	(0.089)	(0.085)
Population	Weight	-0.347	-0.363	-0.205	-0.544
			(0.024)	(0.023)	(0.025)
	Speed	1.043	1.111	1.561	0.599
			(0.062)	(0.057)	(0.061)

Table 7. Estimated Trade-offs in Speed for a 1 lb. Weight Decrease

	Truth	EXP I	EXP II	EXP III
Salespeople	0.667 GHz	0.656 GHz	0.255 GHz	1.64 GHz
Scientists	0.167 GHz	0.150 GHz	0.0420 GHz	0.502 GHz
Population	0.333 GHz	0.326 GHz	0.129 GHz	0.909 GHz

Table 8. Market Share Simulations Using the Estimates from Experiment I

Choice Set	Alternatives	Salespeople		Scientists		Population	
		Truth	Predicted	Truth	Predicted	Truth	Predicted
I	3 lbs, 2.0 GHz	0.667	0.674	0.333	0.317	0.500	0.496
	6 lbs, 3.0 GHz	0.333	0.326	0.667	0.683	0.500	0.504
	3 lbs, 2.0 GHz	0.333	0.401	0.167	0.238	0.250	0.320
II	3 lbs, 2.0 GHz	0.333	0.401	0.167	0.238	0.250	0.320
	6 lbs, 3.0 GHz	0.333	0.197	0.667	0.524	0.500	0.361
	3 lbs, 2.0 GHz	0.667	0.514	0.333	0.191	0.500	0.353
III	6 lbs, 3.0 GHz	0.167	0.243	0.333	0.404	0.250	0.324
	6 lbs, 3.0 GHz	0.167	0.243	0.333	0.404	0.250	0.324

Figure 1. Box Plots of the Individuals' Tastes