



# Substitution Patterns of the Random Coefficients Logit

Thomas J. Steenburgh  
Andrew Ainslie

Working Paper

10-053

Copyright © 2010 by Thomas J. Steenburgh and Andrew Ainslie

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

# Substitution Patterns of the Random Coefficients Logit

Thomas J. Steenburgh and Andrew Ainslie\*

January 11, 2010

## Abstract

Previous research suggests that the random coefficients logit is a highly flexible model that overcomes the problems of the homogeneous logit by allowing for differences in tastes across individuals. The purpose of this paper is to show that this is not true. We prove that the random coefficients logit imposes restrictions on individual choice behavior that limit the types of substitution patterns that can be found through empirical analysis, and we raise fundamental questions about when the model can be used to recover individuals' preferences from their observed choices.

Part of the misunderstanding about the random coefficients logit can be attributed to the lack of cross-level inference in previous research. To overcome this deficiency, we design several Monte Carlo experiments to show what the model predicts at both the individual and the population levels. These experiments show that the random coefficients logit leads a researcher to very different conclusions about individuals' tastes depending on how alternatives are presented in the choice set. In turn, these biased parameter estimates affect counterfactual predictions. In one experiment, the market share predictions for a given alternative in a given choice set range between 17% and 83% depending on how the alternatives are displayed both in the data used for estimation and in the counterfactual scenario under consideration. This occurs even though the market shares observed in the data are always about 50% regardless of the display.

Random Coefficients Logit, Independence from Irrelevant Alternatives, IIA, Rational Choice Theory, Similarity Critique, Heterogeneity, Ecological Fallacy, Cross-level Inference

---

\*The authors thank David E. Bell, Joseph Gavinertvatana, John Gourville, Sunil Gupta, Clarence Lee, and Robert Zeithammer for helpful comments. The authors alone, of course, are responsible for any errors.

# 1 Introduction

Consider a situation in which the same choice set is presented three different ways to a group of decision makers. On the first occasion, the choice set includes two alternatives. Denote this set as  $\{A,B\}$ . On the second occasion, a duplicate for alternative A is added to the choice set, so the set becomes  $\{A,A,B\}$ . On the third occasion, a duplicate for alternative B is added to the original set, so the choice set becomes  $\{A,B,B\}$ . Suppose 60 out of 100 individuals choose alternative A on all three occasions. What do these choices reveal about the individuals' preferences? Observing these data, it would seem reasonable to infer that the individuals' preferences do not depend on how the alternatives are displayed in the choice set. In other words, the individuals have rational preferences. Surprisingly, estimating a random coefficients logit on these data would not lead us to this conclusion.

The purpose of this paper to show that the random coefficients logit is a less flexible model than previous research suggests. We prove that the random coefficients logit imposes restrictions on individual choice behavior that limit the types of substitution patterns that can be found through empirical analysis. It implies that overall demand for an alternative must rise if a perfect substitute for it is added to the choice set. This means that a group of individuals cannot behave rationally across different presentations of a choice set. It also *imposes* disproportionate substitution patterns among market shares rather than recovers them from the data. To help explain our analytical results, we discuss why the population correlations found in the model do not overcome the problems associated with IIA, as previous research suggests [e.g. Berry et al., 1995, Nevo, 2001, Train, 2009], and why the McFadden and Train [2000] theorem has limited implications for empirical analysis.

Part of the misunderstanding about the random coefficients logit can be attributed to the lack of cross-level inference<sup>1</sup> in previous research. To overcome this deficiency, we design several Monte Carlo experiments to show what the model predicts both about individual and population choice behavior. In the first two experiments, we explore what the model predicts when it is estimated on data in which individuals behave rationally, as they are presumed to actually behave. We show that the random coefficients logit results in biased parameter estimates that lead us to very different conclusions about individual preferences depending on how alternatives are presented in the choice set. For example, in one experiment, we infer that an individual would need a 0.26 GHz increase in processor speed to compensate for a 1 lb increase in weight if the model is estimated on data in

---

<sup>1</sup>A researcher conducts cross-level inference by estimating and interpreting quantities at different levels of aggregation. This type of analysis is commonly of interest in the social sciences.

which choices are displayed one way, whereas the same individual would need 2.36 GHz increase in processor speed if they are displayed another. This occurs even though the individual has rational preferences and is observed to make consistent choices across the two presentations.

We show that this finding has implications for policy analysis too, as the bias in parameter estimates compounds upon itself when evaluating counterfactual scenarios. In one experiment, the market share predictions for a given alternative in a given choice set range between 17% and 83% depending on how the alternatives are displayed in the data used for estimation and in the counterfactual scenario under consideration. This occurs even though the market share observed in the data is always about 50% regardless of how the alternatives are displayed.

We conclude by showing what the random coefficients logit leads us to believe when it is estimated on data in which individuals behave with IIA. Paradoxically, we infer that individuals' preferences do not change across presentations in this case. This occurs even though their observed choices suggest that they do.

## 2 The Similarity Critique

### 2.1 Individual Substitution Patterns

Consider the example proposed by Steenburgh [2008] to recall the long-standing argument against models with IIA. An individual faces a choice between two MacBook computers:

	Weight	Processor Speed
MacBook A	3 lbs	2.0 GHz
MacBook B	6 lbs	3.0 GHz

MacBook A is the lighter alternative, but it runs at a slower speed; MacBook B is faster, but heavier. Assume the individual chooses MacBook A with probability  $3/5$ . Now suppose a third alternative, MacBook B', is added to the choice set. MacBook B' is identical to MacBook B, with the prime being used simply for clarity in this example. It weighs 6 lbs and runs at 3.0 GHz.<sup>2</sup> Thus, the individual would be equally likely to choose either MacBook B or MacBook B' if she were asked

---

<sup>2</sup>We will use perfect substitutes throughout the paper because it is clear what the choice probabilities should exactly be in this case. Nevertheless, the basic arguments do not depend on the alternatives being identical and would carry over to close substitutes too.

to choose between just the two of them. What would happen if the individual is presented with a choice among all three alternatives?

The original argument made by Debreu [1960] follows from rational choice theory, which imposes an element of consistency on individual preferences and choice behavior. Rationality presumes that an individual's preference between any two alternatives does not depend on how they are presented in the choice set. For this to occur, MacBook A would need to be chosen with probability  $3/5$ , MacBook B with probability  $1/5$ , and MacBook B' with probability  $1/5$  if the perfect substitute is added to the choice set. This substitution pattern is consistent with rational choice theory because the individual prefers the light MacBook regardless of whether the choice set is displayed as  $\{A,B\}$  or as  $\{A, B, B'\}$ .

The objection to models with IIA is that they do not allow rational substitution to occur. By definition, a model with IIA requires the ratio of any two choice probabilities to remain the same no matter what other alternatives are included in the choice set. If the individual behaves with IIA, MacBook A would be chosen with probability  $3/7$ , MacBook B with probability  $2/7$ , and MacBook B' with probability  $2/7$ . The probability that a fast MacBook is chosen would rise from  $2/5$  in the original set to  $4/7$  in the set that contains a perfect substitute. This new substitution pattern, however, is at odds with rational choice theory because the individual's preference between a light and a fast MacBook now depends on how the alternatives are displayed. The individual prefers the light MacBook when presented with the choice set  $\{A, B\}$ , yet prefers the fast MacBook when presented with the choice set  $\{A, B, B'\}$ . Similar arguments have been made by Savage [Luce and Suppes, 1965], Tversky [1972], and McFadden [1974], among others.

While it is commonly understood that IIA is an undesirable property and we want to use models that do not possess it, it seems to be overlooked that simply breaking IIA is not sufficient to ensure that rational substitution occurs. To clarify this idea, consider the following example. Suppose a third model implies that MacBook A would be chosen with probability  $1/5$ , MacBook B with probability  $2/5$ , and MacBook B' with probability  $2/5$  when MacBook B' is added to the choice set. This new model clearly breaks IIA. But again, the individual prefers the light MacBook when presented with the choice set  $\{A, B\}$ , yet prefers the fast MacBook when presented with the choice set  $\{A, B, B'\}$ . In fact, this model is even more objectionable than the one with IIA because an even smaller proportion of the entrant's choice probability (0% as opposed to 40%) is drawn from the perfect substitute, and the overall probability that a faster MacBook is chosen rises by an even greater amount (an increase from  $2/5$  to  $4/5$  as opposed to an increase from  $2/5$  to  $4/7$ ) when the perfect

substitute is added to the set.

In short, the goal in developing choice models should not be just to break IIA. Rather, it should be to allow rational substitution to occur.

A few additional points are worth noting. First, it is common to discuss the substitution patterns implied by a model rather than the choice probabilities themselves. This is for convenience, as the substitution patterns merely represent how the choice probabilities change across the two choice sets. We will use the following terms to describe the two aforementioned patterns: An individual exhibits *rational substitution* if 100% of the entrant's choice probability is drawn from its perfect substitute and 0% is drawn from the other alternatives. Although they have not previously been named, it is well accepted that these ratios represent the desired substitution patterns when a perfect substitute is added to the choice set. An individual exhibits *proportional substitution* if the entrant's choice probability is drawn from each of the original alternatives in proportion to their choice probabilities. In the previous example, when the individual behaves with IIA and exhibits proportional substitution, 60% of the entrant's choice probability is drawn from MacBook A and 40% is drawn from MacBook B.

Second, although we refer to the desired substitution patterns as rational, there is little disagreement between rational and behavioral economists on this point. For example, in the behavioral economics literature, Tversky [1972] paraphrases the Beethoven / Debussy example proposed in Debreu [1960] to motivate the elimination-by-aspects choice model. Likewise, in the microeconomics literature, the rational substitution patterns are often described as being either more intuitive or more realistic.

## 2.2 Population Substitution Patterns

### 2.2.1 Rational Substitution

Prior research suggests that the random coefficients logit is a highly flexible model that overcomes the problems of the logit because it allows for differences in tastes across individuals [e.g. Berry et al., 1995, Nevo, 2001, Train, 2009]. It has been suggested that the random coefficients logit allows more realistic substitution patterns to be found among market shares, and some prior research even conjectures that the random coefficients logit can recover any substitution pattern that exists in the data [e.g. Train, 1999, Nevo, 2001, Train, 2009]. Although the random coefficients logit implies that individuals behave with IIA, perhaps this is not a concern so long as the researcher's interest centers around only the population's choice behavior.

We show that the random coefficients logit model is less flexible than previous research suggests. First, it implies that overall demand for an alternative must rise if a perfect substitute for it is added to the choice set. Demand rises for every individual in the population, and therefore it must rise for the population as a whole. This means that a population of decision makers cannot exhibit the desired, rational choice behavior even if tastes do vary across individuals. Second, the random coefficients logit implies that disproportionately greater substitution must occur regardless of how individuals truly behave in the data. This means that the random coefficients logit is not recovering more realistic substitution patterns from the observed data; quite to the contrary, it is imposing them.

The following example helps illustrate this point. Suppose there are two types of individuals in the population, Salespeople and Scientists, and the proportion of individuals of each type is  $1/2$ . Both Salespeople and Scientists prefer MacBooks 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 between MacBooks A and B, a Salesperson chooses MacBook A with probability  $2/3$  whereas a Scientist chooses MacBook A with probability  $1/3$ .

The market share of each MacBook 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 MacBook. Thus, the market share of MacBook A is  $\Pr\{A\} = \Pr\{A|Salesperson\}\Pr\{Salesperson\} + \Pr\{A|Scientist\}\Pr\{Scientist\}$ . Taken together, these assumptions imply the following choice probabilities:

	Salespeople	Scientists	Population
MacBook A	$2/3$	$1/3$	$1/2$
MacBook B	$1/3$	$2/3$	$1/2$

Again suppose a third alternative identical to MacBook B, weighing 6.0 lbs and running at 3.0 GHz, is added to the choice set. If the individuals were to exhibit rational substitution, the following choice probabilities and market shares would occur:

	Salespeople	Scientists	Population
MacBook A	$2/3$	$1/3$	$1/2$
MacBook B	$1/6$	$1/3$	$1/4$
MacBook B'	$1/6$	$1/3$	$1/4$

Yet, if both Salespeople and Scientists were to behave with IIA, the following would occur:

	Salespeople	Scientists	Population
MacBook A	1/2	1/5	14/40
MacBook B	1/4	2/5	13/40
MacBook 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 MacBook A,  $1/3$  for MacBook B, and  $1/3$  for MacBook B'. Nevertheless, it should also be clear that allowing for differences in tastes does not resolve the problems created by individuals behaving with IIA. The population does not behave rationally because the market share of the faster MacBook increases from  $1/2$  in the original set to  $13/20$  (or 65%) in the set with a perfect substitute. Expressed in terms of substitution patterns, only 54% of the demand for the entrant is drawn from its perfect substitute, whereas we would expect 100% to be drawn.

The following theorem asserts that this is true in general:

**Theorem 1.** *If individuals behave with IIA, overall demand for an 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  $k_n P_n + k_n P_n + k_n (1 - P_n) = 1$ , where  $k_n$  is a constant that rescales the original choice probabilities. This 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.  $\frac{1}{N} \sum_n 2k_n P_n > \frac{1}{N} \sum_n P_n$  □



At first blush, this theorem seems to be at odds with Theorem 1 of McFadden and Train [2000, p. 451], which shows that a mixed logit can approximate any random utility model to any desired degree of accuracy. As will be shown in section 3, however, the theorem in McFadden and Train has more limited implications than previous research suggests.

### 2.2.2 Disproportionately Greater Substitution

Even though the population does not exhibit rational substitution in the example, it is interesting to note that a disproportionately greater percentage of the entrant's market share is drawn from its pre-existing substitute. We may wonder whether this is always the case with the random coefficients logit. If so, observing this pattern in empirical analysis is not evidence of the random coefficients logit's ability to recover more realistic substitution patterns that exist in the data. The following theorem asserts that this is true in general:

**Theorem 2.** *Consider a population of decision makers who behave with IIA and whose tastes vary across individuals. An new alternative must draw a greater proportion of its market share from its perfect substitute, even though each individual in the population exhibits proportional substitution.*

*Proof.* It suffices to consider a choice between two alternatives. Suppose individual  $n$  chooses one alternative with probability  $P_n$  and a composite of the 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 choice probability of the first alternative changes by  $P_n - k_n P_n = (1 - k_n) P_n$ , where  $k_n = 1/(1 + P_n)$  as before. (IIA implies that each individual exhibits proportional substitution, so for each individual the proportion of the entrant's choice probability that is drawn from its perfect substitute is  $(1 - k_n) P_n / k_n P_n = P_n$ .)

The market share of the first good in the original choice set is  $S = \frac{1}{N} \sum_{\forall n} P_n$ , the change in its market share between choice sets is  $\Delta = \frac{1}{N} \sum_{\forall n} (1 - k_n) P_n$ , and the market share of each perfect substitute in the expanded choice set is  $S' = \frac{1}{N} \sum_{\forall n} k_n P_n$ . We want to show that the market share of the entrant is disproportionately drawn from its perfect substitute:  $\Delta/S' > S$ , which occurs if 
$$\sum_{\forall n} (k_n P_n) P_n > \left( \sum_{\forall n} k_n P_n \right) \left( \frac{1}{N} \sum_{\forall n} P_n \right).$$

If we order the individuals such that  $P_1 \geq P_2 \geq \dots \geq P_n$ , then  $k_1 P_1 \geq k_2 P_2 \geq \dots \geq k_n P_n$ . Thus, Chebyshev's sum inequality implies that 
$$\sum_{\forall n} (k_n P_n) P_n \geq \left( \sum_{\forall n} k_n P_n \right) \left( \frac{1}{N} \sum_{\forall n} P_n \right).$$

If the population has heterogeneous tastes, such that  $P_1 > P_n$ , then  $\sum_{\forall n} (k_n P_n) P_n > \left( \sum_{\forall n} k_n P_n \right) \left( \frac{1}{N} \sum_{\forall n} P_n \right)$ , which completes the proof.

Notice too, however, that if the population has homogeneous tastes, such that  $P_1 = P_2 = \dots = P_n = P$  as the logit model suggests, then  $\sum_{\forall n} (k_n P_n) P_n = \left( \sum_{\forall n} k_n P_n \right) \left( \frac{1}{N} \sum_{\forall n} P_n \right)$ , which implies that the population exhibits proportional substitution because  $\frac{1}{N} \sum_{\forall n} P_n = P$ .  $\square$

The proof can be understood intuitively as follows:  $k_n P_n$  is the probability that individual  $n$  chooses the entrant, and  $P_n$  is the proportion of this probability drawn from the perfect substitute. Thus, the individuals who are most likely to choose the entrant are also the individuals who draw the greatest proportion of demand from the perfect substitute. Chebyshev's sum inequality states that the demand drawn from the perfect substitute under this arrangement,  $\sum_{\forall n} (k_n P_n) P_n$ , is at least as great as the demand that would have been drawn from the perfect substitute if every individual in the population were to substitute away by the average amount,  $\left( \sum_{\forall n} k_n P_n \right) \left( \frac{1}{N} \sum_{\forall n} P_n \right)$ .

Returning to the example, suppose the population consisted of one hundred Salespeople and one hundred Scientists. The demand for MacBook B' would be  $(100 \cdot 1/4) + (100 \cdot 2/5) = 65$  units in the expanded choice set. Of this demand,  $(100 \cdot 1/4) (1/3) + (100 \cdot 2/5) (2/3) = 35$  units would be drawn from MacBook B. This is greater than what would have occurred if the demand for MacBook B' had been drawn in proportion to the original market shares, which would have been  $65 (1/2) = 32.5$  units. This occurs because Scientists are more likely to choose the entrant than Salespeople (40 units vs 25 units) and a greater proportion of the Scientists' demand is drawn from the pre-existing perfect substitute ( $2/3$  vs  $1/3$ ). But obviously, not all of the demand for MacBook B' has been drawn from MacBook B; only  $35/65 \approx 54\%$  it has been.

From one perspective, Theorem 2 seems to suggest that the random coefficients logit represents a step in the right direction. We may object to the fact that each individual exhibits proportional substitution, but the population as a whole is guaranteed to substitute among the alternatives in a way that is more in line with our expectations about how it should behave. Perhaps the model is adequate if interest centers around only the population's choice behavior. We will take up this issue through Monte Carlo simulations in section 4.

Nevertheless, as will be discussed further in the next section, Theorem 2 is a bit unsettling in its resemblance to the ecological fallacy, which describes situations in which the behavior of an individual and the behavior of a population do not correspond [Robinson, 1950]. The theorem implies that the substitution behavior of each individual *must* differ from that of the population. Furthermore,

prior research suggests that the flexibility of the random coefficients logit model allows it to capture more realistic substitution patterns that occur in the data. As it turns out, the opposite is true. The random coefficients logit *imposes* disproportionately greater substitution on the population regardless of the observed choice behavior. These patterns are a mathematical consequence of the model.

### 3 The Random Coefficient Logit

In this section, we derive the standard random coefficients logit model, draw a parallel with the earlier work of Robinson [1950] to discuss the relationship between the population correlations and the individual choice behavior, and discuss the implications of the McFadden and Train [2000] theorem. The broader purpose is to help explain why prior work has come to different conclusions about the substitution patterns of the random coefficients logit.

#### 3.1 The Model

The random coefficients logit generalizes the homogeneous logit by allowing tastes to vary across individuals. The term “random” coefficients is misleading in the sense that it does not imply that an individual’s tastes randomly fluctuate across occasions. Nor does it imply anything about parameter uncertainty.<sup>3</sup> Rather, the term simply implies that tastes are modeled as varying across individuals in the population.

The first step in building a random coefficients logit is to define the relationship between an individual’s observed choices and his or her preferences. An individual’s preferences are represented with an additively separable utility function that is decomposed into two components. The utility of individual  $n$  on occasion  $t$  from alternative  $j$  is specified as

$$u_{ntj} = x_{tj}\beta_n + \varepsilon_{ntj} \tag{1}$$

$$\varepsilon_{ntj} \sim EV(I) \text{ for } n = 1, \dots, N; t = 1, \dots, T; j = 1, \dots, J$$

The first component,  $x_{tj}\beta_n$ , is referred to as the observed utility. It is a function of the observed attributes of each alternative on each occasion,  $x_{tj}$ , and the individual’s tastes for those attributes,

---

<sup>3</sup>The terms “random coefficients” and “random effects” are inconsistently used in the literature and often lead to confusion. Gelman and Hill [2007, p. 245] identify five different uses of the terms random and fixed effects and suggest dropping these terms in favor of modeled vs unmodeled.

$\beta_n$ . The second component,  $\varepsilon_{ntj}$ , is referred to as the unobserved utility. It is a random variable that accounts for factors other than the observed attributes that affect the individual's utility.<sup>4</sup>

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

$$Cov(\varepsilon_{ntj}, \varepsilon_{ntk}) = 0 \quad \forall j \neq k \quad (2)$$

The link between the individual's utility and observed choice behavior is established by assuming that the individual chooses the alternative that provides the greatest utility. The decision rule governing his or her behavior is to choose alternative  $j$  on a given occasion if and only if  $u_{ntj} > u_{ntk} \quad \forall k \neq j$ . Conditional on the observed attributes and the individual's tastes, the probability that individual  $n$  chooses alternative  $j$  on occasion  $t$  is

$$\begin{aligned} p(y_{ntj} = 1 \mid x_t, \beta_n) &= \Pr \{ \varepsilon_{ntk} - \varepsilon_{ntj} < x_{tj}\beta_n - x_{tk}\beta_n \quad \forall k \neq j \mid x_t, \beta_n \} \\ &= \frac{\exp(x_{tj}\beta_n)}{\sum_{\forall k} \exp(x_{tk}\beta_n)} \end{aligned} \quad (3)$$

where  $y_{ntj}$  is an indicator variable that takes the value 1 if the alternative is chosen and the value zero otherwise. These are the well-known logit choice probabilities, which imply that each individual behaves with IIA.

Putting these terms together, we can construct the multinomial logit probability mass function

$$p(y_{nt} \mid x_t, \beta_n) = \left( \frac{\exp(x_{t1}\beta_n)}{\sum_{\forall k} \exp(x_{tk}\beta_n)} \right)^{y_{nt1}} * \dots * \left( \frac{\exp(x_{tJ}\beta_n)}{\sum_{\forall k} \exp(x_{tk}\beta_n)} \right)^{y_{ntJ}}$$

where  $y_{nt}$  is a vector whose elements are  $y_{ntj}$ . This is the basic building block of all subsequent models, so to simplify notation from this point forward we will write

$$y_{nt} \sim MNL(x_t\beta_n) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T \quad (4)$$

when the observed choices are generated according to this distribution.

---

<sup>4</sup>Implicit in this specification, the unobserved utilities are assumed to be independent of the observed attributes,  $\varepsilon_{ntj} \perp x_{tk} \quad \forall j, k$ .

The next step is to model how tastes vary across individuals in the population. In the case of a random coefficient logit, the tastes are commonly assumed to follow a multivariate normal distribution and the model is specified as

$$\begin{aligned} y_{nt} &\sim MNL(x_t \beta_n) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T \\ \beta_n &\sim MVN(\theta, \Sigma) \quad \text{for } n = 1, \dots, N \end{aligned} \tag{5}$$

The same model can be written in mixed-effects form 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 becomes

$$\begin{aligned} y_{nt} &\sim MNL(x_t(\theta + \delta_n)) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T \\ \delta_n &\sim MVN(0, \Sigma) \quad \text{for } n = 1, \dots, N \end{aligned} \tag{6}$$

The mixed-effects form is popular in the social sciences and can be useful for analysis [Gelman and Hill, 2007, pp. 264-265], but it should be clear that equations (5) and (6) are equivalent expressions.

In contrast to the random coefficients logit, the homogeneous logit assumes that the tastes of every individual in the population are the same. The homogeneous logit requires a single level and is written as

$$y_{nt} \sim MNL(x_t \beta) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T \tag{7}$$

It is analogous to classical regression model in the sense that the observed choices can be pooled across individuals when estimating tastes.

### 3.2 Population (Ecological) Correlations and Individual Choice Behavior

Prior research [e.g. Berry et al., 1995, Nevo, 2001, Train, 2009] suggests that the problem with the homogeneous logit is that the random component of utility, which includes just  $\varepsilon$ , is both additively separable from the observed utility and is independent and identically distributed. It argues that the lack of correlation gives rise to the IIA property and its restrictive substitution patterns, whereas the random coefficients logit overcomes this issue because it possesses a random component of utility that is no longer independent of the observed product attributes and can change depending on the similarity of the alternatives in the choice set. The purpose of this section is to make clear that these population correlations do not imply that the problems with IIA have been solved.

Train [2009, p. 139] shows that population (ecological) correlations exist in the random coefficients logit as follows. Beginning with the mixed effects specification in equation (6), a composite component of random utility is formed by defining  $\eta_{ntj} \equiv x_{tj}\delta_n + \varepsilon_{ntj}$ . This quantity treats the variation in tastes *across* individuals as being random in addition to the unobserved utility. Correlations between different alternatives exist in the composite random utility, as the covariance between  $\eta_{ntj}$  and  $\eta_{ntk}$  is

$$\begin{aligned}\sigma_{.tjk} &= Cov(\eta_{ntj}, \eta_{ntk}) \\ &= x'_{tj}\Sigma x_{tk} \quad \forall j \neq k\end{aligned}\tag{8}$$

$\sigma_{.tjk}$  is a function of  $\Sigma$  because it is determined by integrating over the distribution of tastes in the population, and therefore it is a property of the population as opposed to a property of an individual. Correlations between different alternatives exist even if  $\Sigma$  is a diagonal matrix, which it is often assumed to be.

Although correlations now exist at the population level of the model, this does not mean that these are the right correlations. The ecological fallacy should make us cautious about assuming that these correlations correct any problems at the individual-level. Robinson [1950], the study that made this fallacy well known, shows that individual- and population-level correlations can bear little resemblance to one another. An analogous relationship exists here, as the random utilities are correlated at the population-level of the model (equation 8), yet are uncorrelated by assumption at the individual-level (equation 2).<sup>5</sup> So why should we believe that the presence of population correlations solves the problems with IIA, an individual-level property of the model?

One reason is that some prior research wrongly determines the individual choice probabilities. For example, Train [2009, Ch. 6] asserts that the individual choice probabilities in a random coefficient logit would be

$$P_{ntj} = \int \left( \frac{\exp(x_{tj}\beta)}{\sum_{\forall k} \exp(x_{tk}\beta)} \right) MVN(\beta | \theta, \Sigma) d\beta\tag{9}$$

---

<sup>5</sup>In contrast to these assumptions, a recent working paper by Dotson et al. [2009] allows the random component of utility to depend on the observed attributes at the individual-level of the model. This may be a useful direction for future research.

for the model in equation (5). It makes the argument (pp. 137-139) that the individual choice probabilities in equation (3) are conditional on  $\beta_n$ , which is unobserved by the researcher. If the researcher observed  $\beta_n$ , then the choice probabilities of the random coefficients logit would be same as those of the homogeneous logit, since the  $\varepsilon_{ntj}$  are iid extreme value. But the researcher does not know  $\beta_n$ , and thus must integrate over all possible values that it may take.

This argument, however, is incorrect because it confuses the concept of parameter uncertainty with the concept of taste variation. The quantity in equation (9) is not the probability that individual  $n$  chooses alternative  $j$  on occasion  $t$ . While it is true that the researcher does not observe any individual decision maker's tastes ( $\beta_n$ ) in the random coefficients logit, it is also true that the researcher does not observe any individual's tastes in the homogeneous logit. If this argument were true, the homogeneous logit would not possess IIA either. In fact, not all of the parameter uncertainty has been accounted in equation (9) because the probability is still conditional on the population level parameters  $\theta$  and  $\Sigma$ . As it turns out, the individual choice probability is correctly stated in equation (3), whereas the quantity in equation (9) is akin to a market share.

Another reason for this belief is that many studies never quantify the individual choice probabilities because they either cannot or do not conduct cross-level inference. For example, Berry et al. [1995] and Nevo [2001] proceed by estimating the random coefficients logit on market share data, which makes it impossible to draw individual-level inference. But even when individual-level data are used to estimate the model, the individual-level parameters are often treated as nuisance variables and disregarded. For example, Revelt and Train [1998] and Brownstone and Train [1999] proceed by evaluating the likelihood function

$$\mathcal{L}(\theta, \Sigma) = \int \left[ \prod_{n=1}^N \prod_{t=1}^T MNL(x_{nt}\beta_n) \right] \prod_{n=1}^N MVN(\beta_n | \theta, \Sigma) d\beta_n \quad (10)$$

This approach to estimation makes it impossible to make statements about the individual choice probabilities because the individual decision maker's tastes ( $\beta_n$ ) are integrated out of the likelihood.

The substantive focus of previous research partially explains the lack of cross-level inference. The quantity of interest in these and many other studies is the market share, a population-level quantity. Nevertheless, as we will show through Monte Carlo experiments, assumptions about individual choice behavior can matter in inference even if the ultimate objective is to learn about population-level quantities. We make this clear by taking a fully Bayesian approach to estimation. For example, to estimate the model in equation (5), we simulate from the joint posterior density

$$p(\beta, \theta, \Sigma | y) \propto \left[ \prod_{n=1}^N \prod_{t=1}^T MNL(x_{nt}, \beta_n) \right] \left[ \prod_{n=1}^N MVN(\beta_n | \theta, \Sigma) \right] p(\theta, \Sigma) \quad (11)$$

and draw inference about both the individual-level parameters ( $\beta$ ) and population-level parameters ( $\theta$  and  $\Sigma$ ). This enables us to draw conclusions about the individual choice probabilities as well as the population market shares.

### 3.3 Implications of the McFadden-Train Theorem

The McFadden and Train [2000] theorem is often used to justify the specification of a random coefficient logit [Train, 1998, Revelt and Train, 1998, Nevo, 2000, 2001, Chintagunta et al., 2003, Erdem et al., 2008]. Although it is commonly agreed that the theorem proves that a random coefficients logit is an extremely flexible model, some disagreement does seem to exist about what the theorem specifically shows. Sometimes it is argued that the theorem shows that a random coefficients logit can approximate any random utility model to any desired degree of accuracy [e.g. Train, 1998]. Other times it is argued that the theorem shows that the model allows for a flexible pattern of substitution at the aggregate level, but it imposes IIA at the individual level [e.g. Erdem et al., 2008]. Train [1999, 2009] provide an intuitive explanation of the proof.

The purpose of this section is to clarify what McFadden and Train [2000] theorem means and to show that its implications are more limited for empirical analysis than previous research suggests. We begin by noting that although no distinction seems to be made between a mixed and a random coefficients logit in the literature, the theorem encompasses a broader class of models than just the random coefficients logit. Specifically, the theorem defines any model that can be written in the form

$$p(y_{nj} = 1 | x, \theta) = \int \left( \frac{\exp(x_j \beta)}{\sum_{\forall k} \exp(x_k \beta)} \right) f(\beta | \theta) d\beta \quad (12)$$

as a mixed logit. The theorem formally considers only a single choice occasion, so we will drop the  $t$  subscript.

Following the discussion in Train [1999, p. 128], we will begin by assuming that all individuals have the same tastes. While this may seem to be a strange starting point, the theorem is meant to encompass models of this form too. Not only does this assumption explicitly make the point



that finding the right mixing distribution requires more than finding the right distribution of tastes across individuals, but it also avoids any confusion that might arise about the difference between individual choice probabilities and population market shares, as discussed in the previous section.

Consider the following example. Suppose the true random utility model is

$$\begin{aligned} u_{nj} &= x_j \alpha + \eta_{nj} \\ \eta_{nj} &\sim \text{unif}(-\phi, \phi) \text{ for } n = 1, \dots, N; j = 1, \dots, J \end{aligned} \quad (13)$$

where  $x_j$  represents the observed attributes of alternative  $j$ ,  $\alpha$  is a constant representing an individual's tastes, and  $\eta_{nj}$  is the unobserved utility.

Conditional on  $\alpha$  and  $\eta_{nj}$ , the individual's choices would be deterministic.

$$p(y_{nj} = 1 \mid \alpha, \eta_n) = I(x_j \alpha + \eta_{nj} > x_k \alpha + \eta_{nk} \quad \forall k \neq j)$$

where  $I(\cdot)$  is the indicator function.

The true choice probability is found by integrating over all possible values that  $\eta$  may take, such that

$$\begin{aligned} Q_{nj} &\equiv p(y_{nj} = 1 \mid \alpha, \phi) \\ &= \int p(y_{nj} = 1 \mid \alpha, \eta) p(\eta \mid \phi) d\eta \\ &= \int I(x_j \alpha + \eta_{nj} > x_k \alpha + \eta_{nk} \quad \forall k \neq j) \prod_{k=1}^J \text{unif}(\eta_k \mid \phi) d\eta_k \end{aligned} \quad (14)$$

The goal of the proof is to find a mixed logit that is able to approximate this choice probability arbitrarily closely. Consider the following transformation. Rescale the original utility function by  $1/\lambda$  and then add an iid extreme value term, such that

$$\begin{aligned}
\tilde{u}_{nj} &= \frac{1}{\lambda} (x_j \alpha + \eta_{nj}) + \varepsilon_{nj} \\
\eta_{nj} &\sim \text{unif}(-\phi, \phi) \\
\varepsilon_{nj} &\sim EV(I)
\end{aligned}$$

The re-scaling of utility does not change the model, but the addition of the extreme value term does. The extreme value term, however, is added to the model because doing so produces choice probabilities of the form

$$P_{nj} = \int \left( \frac{\exp\left(\frac{x_j \alpha + \eta_j}{\lambda}\right)}{\sum_{\forall k} \exp\left(\frac{x_k \alpha + \eta_k}{\lambda}\right)} \right) \prod_{k=1}^J \text{unif}(\eta_k | \phi) d\eta_k \quad (15)$$

which satisfy the McFadden and Train [2000] definition of a mixed logit. The theorem then shows that as  $\lambda \rightarrow 0$ , then  $P_j$  approximates  $Q_j$  arbitrarily closely. It is in this sense that a mixed logit can approximate any random utility model to any desired degree of accuracy.

Although the suggested transformation does allow the choice probabilities of the true model to be expressed in the form of a mixed logit, it is less clear that this is a useful transformation for empirical analysis. The theorem makes no distinction between the randomness of the coefficients and randomness of the errors, as the mixing distribution includes both terms. Thus, not only does the researcher need to choose the right distribution for the random coefficients  $\alpha$  (which we did in this example by implicitly assuming that all individuals have the same tastes), but the researcher also has to choose the right distribution for random errors  $\eta$  too. But if the researcher knew that all individuals had the same tastes and that the errors were distributed according to a uniform distribution, there would be no need to approximate the true random utility model with a mixed logit. The true random utility model could simply be written down. The transformation required in the proof puts the choice probabilities in the right form, but it does not make the problem any easier to solve.

In other words, the flexibility implied by the theorem is simply that any random utility model can be expressed in the form of a mixed logit. It does not show that the random coefficients logit specified in equation (5) can approximate any substitution pattern, which seems to be the rationale given in prior research.

Summarizing our analytical results, we have shown that the random coefficients logit is much less flexible than previous research suggests. Not only does it impose IIA at the individual level of the model, but it also imposes specific substitution patterns at the population level, as shown by Theorems 1 and 2. Part of the confusion surrounding this random coefficients logit stems from the fact many empirical studies have made assumptions about what the model predicts without quantifying behavior through cross-level inference. We now turn our attention to this problem and discuss its implications for counterfactual analysis.

## 4 Monte Carlo Choice Experiments

In this section, we design three Monte Carlo experiments to answer questions about the random coefficients logit. In the first two experiments, we ask, “what does the random coefficients logit predict when individuals behave rationally?” In the first experiment, we simplify the problem by assuming that individuals’ tastes fall into two segments. This makes it easy to quantify what the model predicts about both individual and population choice behavior. In the second, we generalize the problem by assuming that tastes are normally distributed across individuals. We show that the model predicts the same types of substitution patterns in this case, and we discuss the implication of our findings for policy analysis, which is of particular interest because the random coefficients logit is often used in structural models of demand. In the third experiment, we ask, “what does the random coefficients logit predict when individuals behave with IIA?”

We take a Monte Carlo approach to the problem for a few reasons. First, it allows us to know how individuals actually behave when making choices as opposed to having to assume it, as we would in a study of real subjects. Also, we can simulate as many choices for each individual in the population as desired, which means we can come to very accurate estimates of individual tastes.

We design the experiments by elaborating on the example discussed in section 2. A population of decision makers is asked to make a series of choices among different configurations of MacBooks, but these choices are presented under three different conditions. In condition A, individuals are asked to choose between two MacBooks. In condition B, a perfect substitute for the lightweight MacBook is added to the choice set, whereas in condition C, a perfect substitute for the high-speed MacBook is added. We observe individuals making very different choices and find that the model leads us to very different conclusions depending on whether individuals behave rationally or with IIA.

## 4.1 Experiment I: Two Segments of Tastes and Rational Choice Behavior

### 4.1.1 Data Generation

We need to make two basic assumptions in order to generate data in each Monte Carlo experiment. First, we have to decide how preferences for weight and speed vary across individuals in the population. Second, we have to decide whether individuals behave rationally when making decisions or whether they behave with IIA. In this experiment, we assume that the individuals' tastes fall into two segments. This makes it easy to compare what the observed choices suggest about individual preferences with what the model predicts about them. We also assume that individuals behave rationally when making decisions.

Specifically, we begin by assuming that each individual's preferences can be represented by a linear utility function. The Salespeople's tastes for weight and speed are assumed to be  $\{-0.4618, 0.6927\}$  and the Scientist's are  $\{-0.2322, 1.3932\}$ .<sup>6</sup> These tastes imply that Salespeople value lighter weights more than Scientists do. A Salesperson would need a 0.67 GHz increase ( $0.4618/0.6927$ ) in processor speed to compensate for a 1 lb increase in weight, whereas a Scientists would need only a 0.17 GHz increase ( $0.2322/1.3932$ ) in processor speed.

Next, we determine the individual choice probabilities. In condition A, the same choice probabilities would arise no matter whether an individual behaves rationally or with IIA because the data contain only two alternatives. Thus, we determine the choice probabilities by assuming that individuals make decisions according to a multinomial logit process. For example, if a Salesperson faces a decision between the following set of MacBooks  $\{(3 \text{ lbs}, 2 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz})\}$ , we determine the choice probabilities to be  $\{2/3, 1/3\}$ .

In conditions B and C, we need to make an additional behavioral assumption in order to determine the choice probabilities. In this experiment, we assume that individuals behave rationally, so that their choices would reveal a consistent set of preferences across the three conditions. Thus, in condition B, when the Salesperson faces a choice among  $\{(3 \text{ lbs}, 2 \text{ GHz}), (3 \text{ lbs}, 2 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz})\}$ , we determine the choice probabilities to be  $\{1/3, 1/3, 1/3\}$ . Although a second lighter MacBook is presented in the choice set, the chance that a Salesperson chooses a lighter alternative remains  $2/3$ . Table 1 contains a summary of the choice probabilities used to generate data.

Finally, we generate choices. We suppose that the population consists of 100 Salespeople and 100

---

<sup>6</sup>We assume values with four-digits of precision for the individuals' tastes because they result in very simple choice probabilities.

Scientists, and we design the experiment such that each individual faces ten repetitions of nine unique choice sets in each condition. This results in 200 individuals making 90 choices in each of three conditions. This design generates enough data to alleviate concerns about sample size when estimating the individual taste parameters, yet enables us to conveniently summarize the observed choices. Table 2 contains a summary of the observed choices.

Central to this experiment, individuals are assumed to behave rationally, so the observed choices should reveal a consistent set of preferences across the three conditions. For example, choice set (viii) presents a decision between a 3 lbs, 2.0 GHz MacBook and a 4.5 lbs, 3.0 GHz MacBook. The Salespeople’s choices should reveal their indifference between these alternatives. This is observed in the data, as Salespeople choose the lighter MacBook 49.8% of the time (498 times in 1,000 occasions) under condition A, 53.8% under B, and 50.3% under C. On the other hand, the Scientists’ choices should reveal a preference for the faster MacBook in this choice set, and they choose a faster alternative 74.9%, 74.6% and 74.0% of the time across the three conditions.

Since the individuals in this experiment make consistent choices across the conditions, the population as a whole makes consistent choices too. For example, the population’s choices should reveal a preference for the faster MacBook in choice set (viii), and its observed market share is 62.6% (1,251 times in 2,000 occasions), 60.4% and 61.9% across the three conditions. Furthermore, the population’s choices should reveal indifference between a lighter and faster MacBook in choice set (iv), and the market share of the lighter MacBook is 49.9%, 50.7% and 50.5% across the three conditions.

These are ideal data sets on which to test the random coefficients logit. Both the individuals’ and the population’s choices reveal a consistent set of preferences regardless of how the alternatives are displayed, so it seems reasonable to expect that random coefficients logit model would lead us to the same inference about the individuals’ preferences regardless of the data set on which it is estimated.

#### 4.1.2 Model and Results

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

$$y_{nt} \sim MNL(x_t \beta_n) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T$$

$$\beta_n \sim MVN(z_n \Theta, \Sigma) \quad \text{for } n = 1, \dots, N$$

where  $z_n$  represents the vector of observed demographic characteristics of individual  $n$ . This vector

is specified with two elements, an intercept common to all individuals and a job type specific to each individual. We standardize job-type such that it takes the value -0.9975 if the individual is a Salesperson and the value 0.9975 if the individual is a Scientist.<sup>7</sup> This and all subsequent models are estimated using `bayesm`, an open-source software package in the R programming language [Rossi et al., 2005].

The population-level regression results are reported in Table 3. The random coefficients logit is able to recover every individual’s tastes when estimated on Data Set A. This is not surprising in that no difference exists between rational and IIA choice behavior when only two alternatives are displayed in the choice set. Nevertheless, the random coefficients logit leads us astray when estimated on the other data sets. In Data Set B, which includes a perfect substitute for the lightweight alternative, the random coefficients logit underestimates every individual’s taste for weight (average taste of -0.198 in B vs. a true value of -0.347) and overestimates their taste for speed (average taste of 1.54 in B vs. a true value of 1.04). In Data Set C, which includes a perfect substitute for the high-speed alternative, the opposite occurs, as the random coefficients logit underestimates every individual’s taste for speed and overestimates their taste for weight.

We show this bias graphically in Figure 1, which includes four panels. The first panel depicts the true distribution of tastes found in the population. The Salespeople’s tastes are represented with the solid red circle and the Scientists’ with the solid red triangle. The remaining three panels superimpose the estimated tastes of each individual (as summarized by the posterior mean of the individual  $\beta_{ns}$ ) on the true distribution. We represent estimates of individual Salespeople’s tastes with open blue circles and Scientists’ tastes with open blue triangles.

As can be seen in the panel denoted Data Set A, the random coefficients logit is able to recover the tastes of both individual Salespeople and Scientists when estimated on data that display only two alternatives. The figure shows that the population consists of two segments. For both Salespeople and Scientists, the clouds of estimated tastes surround their true values. The same does not occur for the other models. In the panel labeled Data Set B, the clouds of estimated tastes shift to the upper right-hand corner of the panel. They no longer surround the true values because the random coefficients logit underestimates the value that both Salespeople and Scientists place on weight relative to processor speed. The opposite occurs in the panel labeled Data Set C. The clouds shift to the lower left-hand corner of the panel because the random coefficients logit underestimates the value that both Salespeople and Scientists place on processor speed relative to weight.

---

<sup>7</sup>The `bayesm` software requires the hierarchical variables to be standardized.

We quantify the magnitude of the trade-off that each type of individual would be willing to make between weight and speed in Table 4. Recall that a Salesperson would need a 0.67 GHz increase in processor speed to compensate for a 1 lb. increase in weight, whereas a Scientist would only need a 0.17 GHz increase. In Data Set A, the random coefficients logit is able to recover these values, as it estimates that a Salesperson with average tastes would need a 0.66 GHz increase in processor speed, and a Scientist with average tastes would need a 0.15 GHz increase.

The random coefficients logit leads us to the wrong conclusions about the individuals' preferences when estimated on the other data sets. In Data Set B, the random coefficients logit implies that a Salesperson with average tastes would need a 0.26 GHz increase in processor speed to compensate for a 1 lb increase in weight and that a Scientist would need a 0.05 GHz increase. By comparison, in Data Set C, the random coefficients logit estimates that a Salesperson with average tastes would need a 2.36 GHz increase in processor speed to compensate for a 1 lb increase in weight and that a Scientist would need a 0.45 GHz increase.

The troubling aspect of this experiment is that the observed choices imply that the individuals' preferences do not change across the three experiments. So why should we infer through the random coefficients logit that they do?

## 4.2 Experiment II: Normally Distributed Tastes and Rational Choice Behavior

### 4.2.1 Data Generation

In our second experiment, we relax the initial assumptions by supposing that tastes are normally distributed across individuals in the population. Although this assumption makes it more difficult to summarize the choice behavior of different individuals, it is done to address concerns that might arise about the simplicity of a two-segment model or about the assumption of a continuous distribution of tastes when the true distribution is discrete.

Specifically, we begin by drawing the tastes of 200 individuals from a multivariate normal distribution with mean  $\theta$  and covariance  $\Sigma$ , where

$$\theta = \begin{pmatrix} -0.347 \\ 1.04 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} (0.1)^2 & 0 \\ 0 & (0.3)^2 \end{bmatrix} \quad (16)$$

The true tastes of individuals in this population are depicted in the upper left-hand panel of Figure 2. The individual who values lighter MacBooks most would need a 2.09 GHz increase in processor speed to compensate for a 1 lb increase in weight, whereas the individual who values faster MacBooks most would need only a 0.08 GHz increase in processor speed.

Next, we determine the individual choice probabilities. In condition A, we do so by assuming that individuals make choices according to a multinomial logit model. In conditions B and C, we assume the probability that an individual chooses a lighter or a faster MacBook does not depend on how the alternatives are presented in the choice set, just as we did in experiment I. Finally, we simulate choices based on the individual choice probabilities.

Table 5 contains a summary of the observed choices. Since individuals are assumed to behave rationally in this experiment, we would expect the population’s choices to reveal a consistent set of preferences across the three conditions. This is observed in the data. For example, the population’s choices should reveal a preference for the faster MacBook in choice set (viii), and its observed market share is 62.2%, 62.5% and 62.2% across the three choice sets.

#### 4.2.2 Model and Results

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

$$\begin{aligned} y_{nt} &\sim MNL(x_t\beta_n) \quad \text{for } n = 1, \dots, N; t = 1, \dots, T \\ \beta_n &\sim MVN(\theta, \Sigma) \quad \text{for } n = 1, \dots, N \end{aligned} \tag{17}$$

We report the population-level regression results in Table 6. Even though we observe individuals behaving rationally across the three choice sets, we infer through the random coefficients logit that their tastes have changed. In Data Set A, the random coefficients logit is able to recover the population’s tastes for weight and speed. By assumption, an individual with average tastes would need a 0.33 GHz increase in processor speed to compensate for a 1 lb increase in weight. The random coefficients logit implies that an individual with average tastes would need a 0.33 GHz increase.

When the random coefficients logit is estimated on the other data sets, it implies that the individuals’ preferences have changed. In Data Set B, the random coefficients logit underestimates the individual tastes for lighter weights. It implies that an individual with average tastes would need only a 0.13 GHz increase in processor speed to compensate for a 1 lb increase in weight. By comparison, in



Data Set C the random coefficients logit underestimates tastes for faster speeds. It implies that an individual with average tastes would need a 0.91 GHz increase in processor speed to compensate for a 1 lb increase in weight.

We show the individual-level parameter estimate graphically in Figure 2. The first panel depicts the true distribution of tastes found in the population, with each individual's tastes being represented with an open red circle. The remaining panels superimpose the estimated tastes of each individual on the true distribution, with the estimates being represented with an open blue circle. As can be seen the panel labeled Data Set A, the random coefficients logit is able to recover the the population's tastes when it is estimated on the data in which only two alternatives are displayed, as the two clouds overlap.

As happened in the previous experiment, the same does not occur when the random coefficients logit is estimated on the other data sets. In the panel labeled Data Set B, the estimated cloud of tastes shifts to the upper right-hand corner of the graph because the random coefficients logit underestimates the value that individuals place on weight relative to processor speed. The opposite occurs in the panel labeled Data Set C. The estimated cloud shifts to the lower left-hand corner of the graph because the random coefficients logit underestimates the value that individuals place on processor speed relative to weight.

Even though we allow for a continuous distribution of tastes, the same troubling problem occurs in this experiment as did in the last. We observe individuals making a consistent set of choices across the three data sets, so it would seem logical to infer that their preferences have not changed. Yet the random coefficients logit leads us to the opposite conclusion.

### 4.2.3 Implications for Policy Analysis

Prior research [e.g. Brownstone and Train, 1999, pp. 126-127] has qualitatively discussed the substitution patterns implied by the random coefficients logit when similar alternatives are added to the choice set, but a complete quantitative analysis has yet to be completed. In this section, we remedy this issue by simulating the market shares predicted by the random coefficients logit in this choice experiment. It could be the case the the random coefficients logit leads us to reasonable market share predictions even though the taste estimates are biased, but we show that this does not happen. In fact, the problems caused by the biased parameter estimates compound upon themselves in counterfactual analysis.

Consider a situation in which Apple is deciding whether to change the way it displays the MacBook in its retail stores. For the sake of discussion, suppose that Apple currently displays a 3 lb, 2.0 GHz MacBook and a 6 lb, 3.0 GHz MacBook on its shelf, and it is wondering what would happen to the market shares if a perfect substitute for either product were added to the display. Given the observed choice behavior (choice set iv in Table 5), a good model would imply that the overall market shares remain 50%-50% no matter how the alternatives are displayed.

Table 7 summarizes the predicted market shares for the three potential displays. Each model does reasonably well in predicting the market shares for choice sets that are observed in the data. Looking down the diagonal of the the table, the random coefficients logit estimated on Data Set A predicts the market share of the lighter alternative to be 49.6% if the choice set  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed. Likewise, the model estimated on Data Set B predicts the market share of the lighter alternative to be 43.3% if  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed, and the model estimated on Data Set C predicts it to be 57.1% if  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed.

Nevertheless, each model produces very skewed predictions when applied to choice sets that are not observed in the data. For example, when estimated on Data Set A, the random coefficients logit predicts the market share of the lighter alternative to be 65.2% if  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed and 34.1% if  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed.

The random coefficients logit skews the market share prediction even further when estimated on the other data sets. For example, when estimated on Data Set B, the random coefficients logit predicts the market share of the lighter alternative to be 17.2% if the choice set  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed. This occurs because the choice set considered in the counterfactual situation is even more different than the one observed in the data. A similar statement can be made about the model estimated on Data Set C. It predicts the market share of the lighter alternative to be 83.2% if  $\{(3 \text{ lbs, } 2.0 \text{ GHz}), (3 \text{ lbs, } 2.0 \text{ GHz}), (6 \text{ lbs, } 3.0 \text{ GHz})\}$  is displayed.

#### 4.2.4 Substitution Patterns

We may wonder why these implications have not been noticed before. One reason is that substantive interest often centers around population-level results, and the random coefficients logit does produce

population-level results that more closely match our intuition than the homogeneous logit does. As Theorem 2 asserts, the random coefficients logit must result in disproportionate substitution among market shares. Nevertheless, this does not imply that the population-level results are correct.

To quantify this idea, we compare the substitution patterns predicted by the random coefficients logit against two standards: the rational substitution patterns that were assumed to create the data and the proportional substitution patterns that would be predicted by a logit model. The statistic that we use to quantify the difference between these patterns is the following substitution ratio:

$$\Psi = \frac{S_{existing|original} - S_{existing|expanded}}{S_{incoming|expanded}}$$

where  $S_{existing|original}$  is the market share of the pre-existing substitute in the original choice set,  $S_{existing|expanded}$  is the share of the pre-existing substitute in the expanded choice set, and  $S_{incoming|expanded}$  is the share of the incoming alternative in the expanded choice set. This statistic is analogous to the substitution ratio proposed in Steenburgh [2008]. It answers the question, “what proportion of the demand for the incoming alternative has been drawn from its pre-existing perfect substitute?”

Table 8 summarizes the substitution ratios. The left-hand side of the table considers what happens when a lightweight substitute is added to the two alternative choice set (Choice Set 1 becomes Choice Set 2 in Table 7.) The right-hand side considers what happens when a high-speed substitute is added (Choice Set 1 becomes Choice Set 3).

We calculate the value of the three ratios as follows: If the population behaves rationally, all of the demand for the incoming MacBook would be drawn from its pre-existing substitute. Thus, rational substitution would imply a ratio of 100%. This is what the model should predict given how individuals actually behave in the observed data.

If the population were to exhibit proportional substitution, the substitution ratio would simply be the market share of the perfect substitute in the two-alternative choice set. For example, working with the market shares in Table 7, if the random coefficients logit is estimated on Data Set A, the market share of the lighter MacBook is predicted to be 49.6% in Choice Set 1. Thus, proportional substitution would imply that 49.6% of the incoming MacBook’s market share is drawn from its pre-existing substitute. This is the standard typically used to compare results.

Finally, we calculate the substitution ratio implied by the random coefficients logit. Continuing with Data Set A, the random coefficients logit predicts that the market share of the lightweight MacBook

is 49.6% in Choice Set 1. The combined market share of the lightweight MacBooks in the expanded choice set is 65.2%, so the market share of each alternative is 32.6%. Thus, the random coefficients logit implies that  $(49.6-32.6)/32.6 = 52.1\%$  of the incoming MacBook’s market share is drawn from its pre-existing substitute.

Table 8 helps explain why the random coefficients logit is thought to be good model. A researcher working with observational data would compare the prediction of the random coefficients logit against the proportional substitution standard. For example, if the random coefficients logit were estimated on Data Set A, it would predict that the incoming lightweight MacBook draws 52.1% of its market share from its perfect substitute instead of only 49.6%. The same would occur for the high-speed MacBook, as the random coefficients logit would predict 52.9% instead of only 50.4%. This would bolster the researcher’s confidence in the model because it confirms the prior belief about what should be happening in the data, at least if the researcher chose to look at close instead of perfect substitutes. As a consequence of Theorem 2, we know that the researcher would always reach this conclusion.

Nevertheless, Table 8 also shows that the random coefficients logit is not a completely flexible model. The model never predicts the actual substitution patterns that exist in the data, as it never results in a substitution ratio of 100%. Furthermore, the predicted substitution patterns, like the predicted market shares, depend on how the alternatives are displayed in the data. When estimated on Data Set B, the random coefficients logit predicts that only 32.0% of the incoming lightweight MacBook’s market share would be drawn from its pre-existing substitute. When estimated on Data Set C, it predicts that 72.8% would be drawn. When the predicted market shares are lower, the substitution ratio is lower too because the random coefficients logit implies that each individual exhibits proportional substitution even if the population as a whole does not.

### **4.3 Experiment III: Normally Distributed Tastes and IIA Choice Behavior**

If the random coefficients logit cannot recover rational substitution patterns, we may wonder what it is able to recover. In this section, we discuss what happens when the random coefficients logit is estimated on data in which individuals behave with IIA.

### 4.3.1 Data Generation

In the final experiment, we keep the assumption that tastes are normally distributed across individuals, but we modify the procedure by assuming that individuals behave with IIA instead of behaving rationally. Specifically, we begin by assuming that individuals have the same tastes as they do in the previous experiment.

Then, we determine the choice probabilities by assuming that each individual makes decisions according to a multinomial logit under every condition. For example, suppose an individual has tastes for weight and speed of  $\{-1/3, 1\}$  and faces the alternatives in choice set (iv). In condition A, when the individual faces a choice between  $\{(3 \text{ lbs}, 2 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz})\}$ , we would determine the choice probabilities to be  $\{1/2, 1/2\}$ . In condition B, when the individual faces a decision among  $\{(3 \text{ lbs}, 2 \text{ GHz}), (3 \text{ lbs}, 2 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz})\}$ , we would determine the choice probabilities to be  $\{1/3, 1/3, 1/3\}$ . In condition C, when the individual faces a decision among  $\{(3 \text{ lbs}, 2 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz}), (6 \text{ lbs}, 3.0 \text{ GHz})\}$ , we would determine the choice probabilities to be  $\{1/3, 1/3, 1/3\}$ . In contrast to the previous experiments, the probability of choosing either a lighter or a faster MacBook now depends on how the alternatives are presented in the choice set. The individual would choose a lighter MacBook with probabilities  $1/2$ ,  $2/3$ , and  $1/3$  across the three experiments.

Finally, we simulate choices based on the individual choice probabilities. These choices are summarized in Table 9. The population's choices suggest that individuals' have context-dependent preferences. This is not surprising given that individuals are assumed to behave with IIA. Consider choice set (iv). In condition A, the market share of the lighter alternative is 49.9%, so the observed choices suggest the population is indifferent between a lighter and a faster MacBook. In condition B, its market share is 65.9%, so the observed choices suggest the population prefers a lighter MacBook. In condition C, its market share is 33.5%, so the observed choices suggest the population prefers a faster MacBook.

### 4.3.2 Model and Results

To infer the individuals' tastes from their observed choices, we estimate the random coefficients logit specified in equation (17). We report the population-level regression results in Table 10. In contrast to the previous two experiments, the random coefficients logit model is able to recover the assumed parameters under every condition. From this we would infer that the individuals' preferences are the same across the three conditions. The model implies that an individual with average tastes

would need a 0.33 GHz increase in processor speed to compensate for a 1 lb increase in weight under condition A, a 0.34 GHz increase under B, and a 0.33 GHz increase under C. This is in stark contrast to the preferences found in experiment II.

We show the individual-level parameter estimate graphically in Figure 3. The first panel depicts the true distribution of tastes found in the population, with each individuals' tastes being represented with an open red circle. The remaining panels superimpose the estimated tastes of each individual on the true distribution, with the estimates being represented with an open blue circle. Across all three conditions, the clouds of estimated tastes no longer shift across the conditions, but rather the surround the cloud of true tastes as they should. Taken at face value these results seem reassuring, as the random coefficients logit is able to recover the correct parameter estimates if individuals do behave with IIA.

On the other hand, our experimental results lead to the following paradox: If individuals actually do behave rationally, the data will suggest that their preferences remain constant across the three conditions. Yet, when we estimate the random coefficients logit on these data, we will infer that their tastes have changed. In experiment II, the population chooses a lightweight MacBook 9,080, 9,044 and 9,053 times out of 18,000 occasions across the three conditions, yet the model implies that an individual with average tastes would need a 0.33, 0.13 and 0.91 GHz increase in processor speed to compensate for a 1 lb increase in weight.

Conversely, if individuals actually do behave with IIA, the data will suggest that their preferences are context-dependent. Yet, when we estimate the random coefficients logit on these data, we will infer that their tastes remain the same. In experiment III, the population chooses a lightweight MacBook 8,896, 11,853 and 6,120 times out of 18,000 occasions across the three conditions, yet the model implies that an individual with average tastes would need a 0.33, 0.34 and 0.33 GHz increase in processor speed to compensate for a 1 lb increase in weight.

## 5 Conclusion

We show that the random coefficients logit imposes restrictions on individual and population choice behavior. These restrictions do not merely limit the range of substitution patterns that can be recovered in empirical analysis. Rather, they raise basic questions about what it means when we say that the random coefficients logit can recover an individual's preferences from their observed

choices. Why would we infer that individuals' preferences depend on the presentation of the choice set *unless* we observe their choices to be changing across presentations?

Why are these fundamental properties not obvious? One reason is that substantive interest often centers around population-level results, and the random coefficients logit produces population-level results that are more in line with our prior expectations than the homogeneous logit does. Ecological correlations are found in the composite random utilities and disproportional substitution must occur among market shares. Nevertheless, as we show through cross-level inference, this does not mean that the results are correct and that the model can be used to conduct policy experiments.

Another reason is that the nature of the problem has been misunderstood. Many choice models have been proposed with the goal of breaking IIA. But why? Simply breaking IIA, either at the individual or at the population level, does not imply either that the similarity critique [Debreu, 1960] has been addressed or that the model is a good model. The real goal should be to build models that help us understand what observed choices reveal about individual preferences. If we presume that individuals could behave rationally, then the model should at least allow it as a possibility.

A final reason is that the problem is difficult. King [1997] made a significant contribution to the related problem of ecological inference more than seventy-five years after Ogburn and Goltra [1919] identified it and forty-five years after Robinson [1950] made it well known. In the interim, many scholars wrote fruitfully on the topic, while others suggested either that the problem did not exist or that it was folly to try and solve it [King, 1997, p. 6]. Perhaps we should expect there is more to learn.

The heart of the matter is that IIA is an assumption about how an individual substitutes among alternatives. Concerns about it cannot be addressed by allowing for differences in tastes across individuals.

## References

- S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.
- David Brownstone and Kenneth Train. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89(1-2):109–129, November 1999.
- 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.
- Jeffrey Dotson, Peter Lenk, Jeffrey P. Brazell, Thomas Otter, Steven N. MacEachern, and Greg M. Allenby. A probit model with structured covariance for similarity effects and source of volume calculations. Technical report, SSRN, 2009.
- 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.
- Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY, 2007.
- Gary King. *A Solution to the Ecological Inference Problem*. Princeton University Press, 1997.
- 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.
- Aviv Nevo. A practitioner’s guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy*, 9(4):513–548, Winter 2000.
- Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, March 2001.



- William F. Ogburn and Inez Goltra. How women vote. *Political Science Quarterly*, 34(3):413–433, 1919. ISSN 00323195. URL <http://www.jstor.org/stable/2141684>.
- 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.
- 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. Recreation demand models with taste differences over people. *Land Economics*, 74(2):230–239, 1998. ISSN 00237639. URL <http://www.jstor.org/stable/3147053>.
- Kenneth E. Train. Mixed logit models for recreation demand. In Joseph A. Herriges and Catherine L. Kling, editors, *Valuing recreation and the environment: revealed preference methods in theory and practice*, chapter 4, pages 121–140. Edward Elgar, 1999.
- Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, U.K., second edition, 2009.
- Amos Tversky. Elimination by aspects: A theory of choice. *Psychological Review*, 79(4):281–299, 1972.

Table 1: Choice Probabilities Experiment I

Choice Set	Alternatives	Data Set A			Data Set B			Data Set C		
		Salespeople	Scientists	Scientists	Salespeople	Scientists	Scientists	Salespeople	Scientists	Scientists
i	4.5 lbs., 2.0 GHz	0.586	0.414	0.414	0.293, 0.293	0.207, 0.207	0.586	0.586	0.414	
	6.0 lbs, 2.5 GHz	0.414	0.586	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293	0.414	
ii	3.0 lbs., 2.0 GHz	0.739	0.500	0.500	0.369, 0.369	0.250, 0.250	0.739	0.500	0.500	
	6.0 lbs., 2.5 GHz	0.261	0.500	0.500	0.261	0.500	0.131, 0.131	0.250, 0.250	0.500	
iii	4.5 lbs, 2.0 GHz	0.500	0.260	0.260	0.250, 0.250	0.130, 0.130	0.500	0.500	0.260	
	6.0 lbs, 3.0 GHz	0.500	0.740	0.740	0.500	0.740	0.250, 0.250	0.370, 0.370	0.740	
iv	3.0 lbs, 2.0 GHz	0.667	0.333	0.333	0.333, 0.333	0.166, 0.166	0.667	0.667	0.333	
	6.0 lbs, 3.0 GHz	0.333	0.667	0.667	0.333	0.667	0.167, 0.167	0.333, 0.333	0.667	
v	4.5 lbs, 2.5 GHz	0.586	0.414	0.414	0.293, 0.293	0.207, 0.207	0.586	0.586	0.414	
	6.0 lbs, 3.0 GHz	0.414	0.586	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293	0.414	
vi	3.0 lbs, 2.5 GHz	0.739	0.500	0.500	0.369, 0.369	0.250, 0.250	0.739	0.500	0.500	
	6.0 lbs, 3.0 GHz	0.261	0.500	0.500	0.261	0.500	0.131, 0.131	0.250, 0.250	0.500	
vii	3.0 lbs, 2.0 GHz	0.586	0.414	0.414	0.293, 0.293	0.207, 0.207	0.586	0.586	0.414	
	4.5 lbs, 2.5 GHz	0.739	0.586	0.586	0.739	0.586	0.207, 0.207	0.293, 0.293	0.586	
viii	3.0 lbs, 2.0 GHz	0.500	0.260	0.260	0.250, 0.250	0.130, 0.130	0.500	0.500	0.260	
	4.5 lbs, 3.0 GHz	0.500	0.740	0.740	0.500	0.740	0.250, 0.250	0.370, 0.370	0.740	
ix	3.0 lbs, 2.5 GHz	0.586	0.414	0.414	0.293, 0.293	0.207, 0.207	0.586	0.586	0.414	
	4.5 lbs, 3.0 GHz	0.414	0.586	0.586	0.414	0.586	0.207, 0.207	0.293, 0.293	0.414	

Table 2: Observed Choices for Experiment I

Choice Set	Alternatives	Data Set A		Data Set B		Data Set C	
		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 Speed	3472	5597	3423	5451	1701, 1741	2697, 2749

Table 3: Population-level Estimates for Experiment I

Regression Coefficients ( $\Delta$ )		Truth	Data Set A	Data Set B	Data Set C
Weight	Intercept	-0.347	-0.357 (0.019)	-0.198 (0.019)	-0.543 (0.019)
	Job Type	0.115	0.144 (0.017)	0.090 (0.017)	0.136 (0.018)
Speed	Intercept	1.04	1.09 (0.06)	1.54 (0.06)	0.599 (0.053)
	Job Type	0.351	0.33 (0.05)	0.45 (0.05)	0.312 (0.052)

Table 4: Estimated Trade-offs in Speed for a 1 lb. Weight Decrease

	Truth	Data Set A	Data Set B	Data Set C
Salespeople	0.67 GHz	0.66 GHz	0.26 GHz	2.36 GHz
Scientists	0.17	0.15	0.05	0.45

Table 5: Observed Choices for Experiment II

Choice Set	Alternatives	Data Set A	Data Set B	Data Set C
i	4.5 lbs, 2.0 GHz	1010	516, 496	981
	6.0 lbs, 2.5 GHz	990	988	494, 525
ii	3.0 lbs, 2.0 GHz	1277	607, 632	1252
	6.0 lbs, 2.5 GHz	723	761	363, 385
iii	4.5 lbs, 2.0 GHz	757	352, 370	799
	6.0 lbs, 3.0 GHz	1243	1278	627, 574
iv	3.0 lbs, 2.0 GHz	1019	502, 501	1015
	6.0 lbs, 3.0 GHz	981	997	494, 491
v	4.5 lbs, 2.5 GHz	1002	520, 507	1027
	6.0 lbs, 3.0 GHz	998	973	462, 511
vi	3.0 lbs, 2.5 GHz	1205	618, 635	1207
	6.0 lbs, 3.0 GHz	795	747	376, 417
vii	3.0 lbs, 2.0 GHz	1027	537, 477	989
	4.5 lbs, 2.5 GHz	973	986	484, 527
viii	3.0 lbs, 2.0 GHz	756	372, 379	757
	4.5 lbs, 3.0 GHz	1244	1249	632, 611
ix	3.0 lbs, 2.5 GHz	1027	525, 498	1026
	4.5 lbs, 3.0 GHz	973	977	456, 518
Totals	Lighter Weight	9080	4549, 4495	9053
	Faster Speed	8920	8956	4388, 4559

Table 6: Population-level Parameter Estimates for Experiment II

Regression Coefficients ( $\Delta$ )				
	Truth	Data Set A	Data Set B	Data Set C
Weight	-0.347	-0.352 (0.025)	-0.206 (0.024)	-0.528 (0.024)
Speed	1.04	1.07 (0.06)	1.60 (0.06)	0.580 (0.060)

Table 7: Predicted Market Shares

		Data Set A	Data Set B	Data Set C
Choice Set 1	3 lbs, 2.0 GHz	49.6%	28.6%	71.9%
	6 lbs, 3.0 GHz	50.4	71.4	28.1
Choice Set 2	{3 lbs, 2.0 GHz; 3 lbs, 2.0 GHz}	65.2	43.3	83.2
	6 lbs, 3.0 GHz	34.8	56.7	16.8
Choice Set 3	3 lbs, 2.0 GHz	34.1	17.2	57.1
	{6 lbs, 3.0 GHz; 6 lbs, 3.0 GHz}	65.9	82.8	42.9

Table 8: Substitution Ratios

	Lightweight Substitutes			High-Speed Substitutes		
	Data A	Data B	Data C	Data A	Data B	Data C
Rational Substitution	100%	100%	100%	100%	100%	100%
Proportional Substitution	49.6	28.6	71.9	50.4	71.4	28.1
Predicted by Mixed Logit	52.1	32.0	72.8	52.9	72.5	31.1

Table 9: Observed Choices for Experiment III

Choice Set	Alternatives	Data Set A	Data Set B	Data Set C
i	4.5 lbs, 2.0 GHz	986	678, 655	685
	6.0 lbs, 2.5 GHz	1014	667	665, 650
ii	3.0 lbs, 2.0 GHz	1255	751, 793	938
	6.0 lbs., 2.5 GHz	745	456	535, 527
iii	4.5 lbs, 2.0 GHz	750	558, 536	488
	6.0 lbs, 3.0 GHz	1250	906	748, 764
iv	3.0 lbs, 2.0 GHz	978	651, 666	669
	6.0 lbs, 3.0 GHz	1022	683	687, 644
v	4.5 lbs, 2.5 GHz	966	644, 668	649
	6.0 lbs, 3.0 GHz	1034	688	676, 675
vi	3.0 lbs, 2.5 GHz	1262	769, 783	883
	6.0 lbs, 3.0 GHz	738	448	566, 551
vii	3.0 lbs, 2.0 GHz	944	636, 670	691
	4.5 lbs, 2.5 GHz	1056	694	648, 661
viii	3.0 lbs, 2.0 GHz	752	566, 508	458
	4.5 lbs, 3.0 GHz	1248	926	750, 792
ix	3.0 lbs, 2.5 GHz	1003	630, 691	659
	4.5 lbs, 3.0 GHz	997	679	680, 661
Totals	Lighter Weight	8896	5883, 5970	6120
	Faster Speed	9104	6147	5955, 5925

Table 10: Population-level Estimates ( $\Delta$ ) for Experiment III

Regression Coefficients ( $\Delta$ )				
	Truth	Data Set A	Data Set B	Data Set C
Weight	-0.347	-0.361 (0.025)	-0.376 (0.026)	-0.338 (0.025)
Speed	1.04	1.11 (0.06)	1.11 (0.06)	1.03 (0.06)

Figure 1: Distributions of Tastes in Experiment I

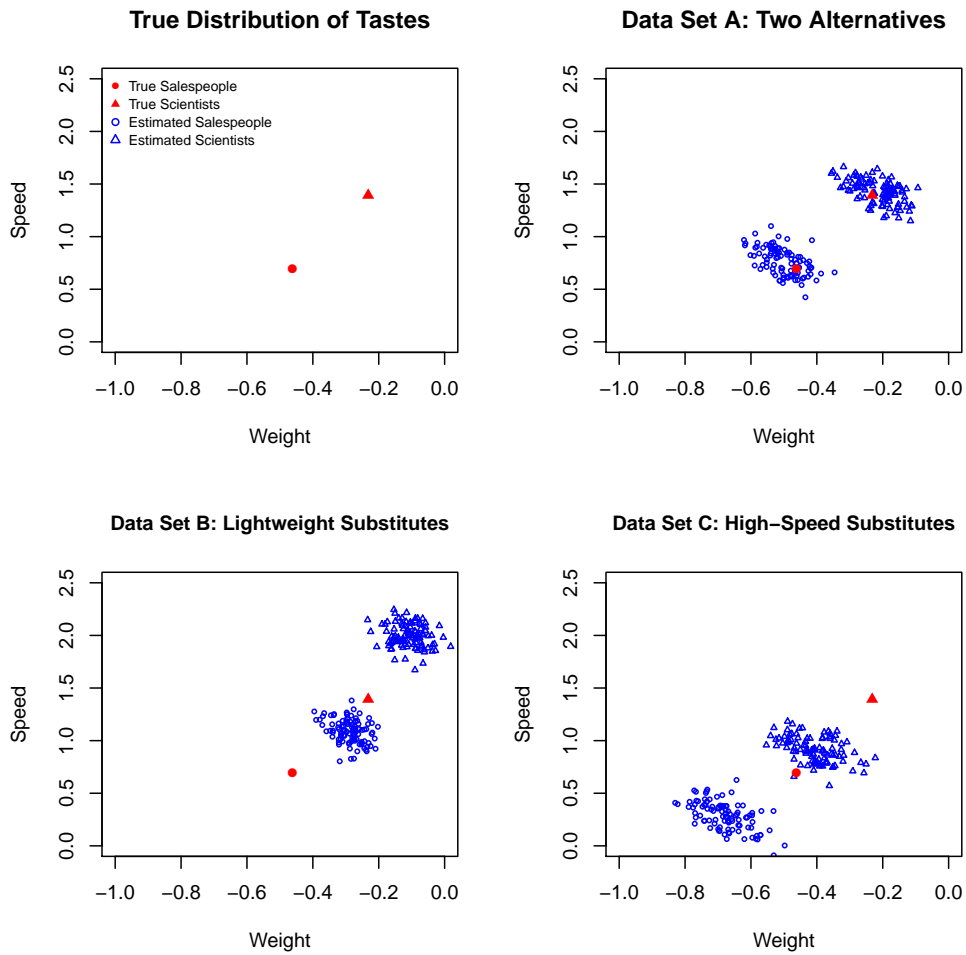


Figure 2: Distributions of Tastes in Experiment II

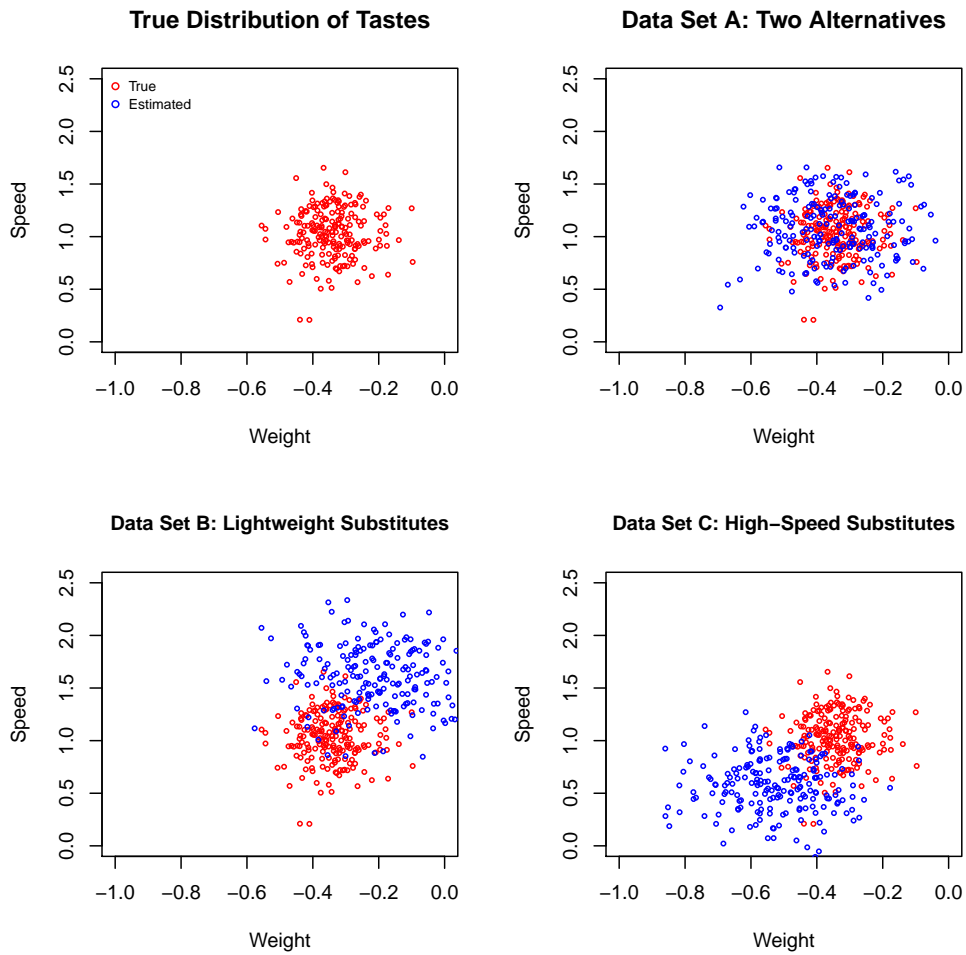




Figure 3: Distributions of Tastes in Experiment III

