The Annals of Statistics
1988, Vol. 16, No. 2, 887–894

TWO-STAGE BANDITS

BY MURRAY K. CLAYTON AND JEFFREY A. WITMER

University of Wisconsin and Oberlin College

Two stochastic processes, or "arms," that yield dichotomous responses are available for use in a two-stage decision problem. During the first stage, arms are chosen sequentially; the resulting observations are discounted by a fixed value β. A single arm must be used in the second stage, in which observations are not discounted. The decision to end the first stage is based on the data obtained. Optimal strategies are considered in the presence of the random discount sequence that arises in this setting. This extends the work of Berry and Fristedt (1979).

1. Introduction. Consider a "two-stage bandit" problem. Assume that there are two "arms" (or machines, treatments, etc.) that yield dichotomous responses —success or failure. The characteristics of the arms are at least partially unknown, and so learning can take place. Observations or "pulls" on the arms may be made in two "stages." In the first stage, pulls can be made on both arms, but in the second stage, all pulls must be made on the same arm.

We consider sequential strategies for this problem. [Nonsequential strategies for the two-stage problem are discussed in Clayton and Witmer (1986) and Witmer (1986)]. In this context, a strategy is a decision rule that specifies, for any history of pulls and their outcomes, which arm to select for the next pull. When the first stage has not yet ended, a strategy must specify whether to continue the first stage. If the first stage is ended, the strategy must specify which arm is to be pulled throughout the second stage. Note that the first-stage length, denoted here by K, may be random. We allow for the possibility that the total number of pulls to be made, N, is not known at the onset.

Let X_i denote the outcome of the ith pull of arm 1; let Y_i denote the corresponding quantity for arm 2. We assume that the sequences $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are independent of each other. The success rate on arm 1, θ_1, is assumed known, whereas the success rate on arm 2, θ_2, is unknown. Information regarding the effectiveness of arm 2 is summarized by a prior distribution, π, on θ_2. Conditional on θ_i, the results on arm i are assumed to be independent; unconditionally, they are exchangeable.

We consider discounting of first-stage successes by a factor β, where $0 \leq \beta \leq 1$, such that each success in the first stage has utility β, whereas a success in the second stage has utility 1. "Optimal" strategies are defined in this setting to be those that maximize the total expected discounted utility. When $\beta = 1$, this setting coincides with that of a one-armed bandit [see Berry and Fristedt (1979, 1985)]. A two-stage problem with a sequential choice of first-stage length has also

Received February 1987; revised July 1987.
AMS 1980 *subject classification*. Primary 62C10.
Key words and phrases. Two-stage bandit, sequential decisions, regular discounting, random discounting.

been discussed by Petkau (1978), wherein first-stage observations are discounted by an additive "cost per observation," rather than by a multiplicative factor, β, as applies to the current setting. In a broader sense, the current model can be viewed as a modification of the two-stage models of Canner (1970) and Witmer (1986). This issue is discussed in further detail in Clayton and Witmer (1986). See Berry and Fristedt (1985) for a discussion of other approaches to two-stage problems.

Ideally, we would give an explicit specification of an optimal strategy for any θ_1, π, β and distribution for N. However, this is next to impossible unless N is known and is small. In most cases the determination of optimal strategies will typically require the use of a computer. Consequently, knowing the properties of such strategies will be helpful in reducing the number of strategies that must be considered in such an explicit determination. Therefore, we give a partial characterization of optimal strategies and of optimal values of K.

Our main result is a generalization of Theorem 2.1 of Berry and Fristedt (1979). We consider the case $\beta < 1$, which stands in contrast to the case $\beta = 1$. This leads to a discussion of random discount sequences.

2. Regular discount sequences. Let τ denote a given strategy, and let τ_i denote the outcome, 0 or 1, for pull i when following τ. When the total number of pulls to be made, N, is known the expected utility of τ is

$$u(\tau|N) = E\left\{ \sum_{i=1}^{K} \beta\tau_i + \sum_{i=K+1}^{N} \tau_i \right\}. \tag{2.1}$$

When N is unknown we suppose that N is a random variable with a probability distribution, Q, on the positive integers and that N is independent of the data X_i, Y_i. In addition, we assume that $E(N) < \infty$. Let $P(N = n|Q) = p_n$ and let $\alpha_n = P(N \geq n|Q)$. Let $\delta_i = \beta\tau_i$ for $i \leq K$ and let $\delta_i = \tau_i$ for $i > K$. Then the expected utility of a strategy τ is

$$\begin{aligned} u(\tau|Q) &= \sum_{n=1}^{\infty} u(\tau|n)P(N = n|Q) \\ &= \sum_{n=1}^{\infty} p_n E\left(\sum_{i=1}^{n} \delta_i|Q \right) \\ &= \sum_{i=1}^{\infty} E\left\{ (\delta_i|Q) \sum_{n=i}^{\infty} p_n \right\} \\ &= \sum_{n=1}^{\infty} E\{(\delta_i|Q)\alpha_i\}. \end{aligned} \tag{2.2}$$

Equation (2.2) lends an intuitive understanding to the role of the discount factors α_i and β: Each potential observation, τ_i, is discounted by the probability that it is actually observed, α_i. First-stage observations are also discounted by the discount factor β. We refer to the sequence $A_\beta = (\beta\alpha_1, \beta\alpha_2, \ldots,$

$\beta\alpha_K, \alpha_{K+1}, \ldots)$ as the "effective" discount sequence; the sequence $A = (\alpha_1, \alpha_2, \ldots)$ we call the "distribution" discount sequence.

An equivalent and more natural expression of (2.2) is

$$u(\tau) = E\left\{ \sum_{i=1}^{K} \beta\tau_i\alpha_i + \sum_{i=K+1}^{\infty} \tau_i\alpha_i \right\}, \tag{2.3}$$

suppressing the dependence on Q.

Intuitively, one might expect that the first stage would be used exclusively for experimentation with the unknown arm. However, the following example shows that it is possible for arm 1 to be optimal in the first stage.

EXAMPLE 2.1. Suppose Q is such that $P(N = 1) = 0.9 = 1 - P(N = 10)$. Furthermore, assume that $\beta \in [0.923, 1]$, $\theta_1 = 0.55$ and that π is a uniform distribution on $[0, 1]$. By evaluating all possible strategies, we can show that it is uniquely optimal to pull arm 1 initially (in the first stage).

We are interested in those cases in which it is optimal to never use arm 1 in the first stage. Berry and Fristedt (1979) have characterized such cases when $\beta = 1$, i.e., for the "standard" one-armed bandit problem. This characterization involves consideration of *regular* distribution discount sequences.

DEFINITION 2.1. A discount sequence $A \equiv (\alpha_1, \alpha_2, \ldots)$ is *regular* if, for each m, $\gamma_m\gamma_{m+2} \leq \gamma_{m+1}^2$, where $\gamma_m \equiv \sum_{i=m}^{\infty}\alpha_i$. A distribution is said to be regular if it yields a regular discount sequence.

Note that the class of regular distributions includes the geometric, discrete uniform, binomial and Poisson distributions on N. Also note that if $(\alpha_1, \alpha_2, \ldots)$ is regular, then so is $(\alpha_i, \alpha_{i+1}, \ldots)$, for all i.

Although Berry and Fristedt did not develop their results explicitly for the two-stage setting under discussion here, it is possible to restate their Theorem 2.1 as follows.

THEOREM 2.1 [Berry and Fristedt (1979)]. *For $\beta = 1$, for all π and for all $\theta_1 \in [0, 1]$, there exists an optimal strategy under which arm 1 is never pulled in the first stage if and only if the distribution discount sequence A is regular.*

Theorem 2.2 extends this result to the general case of $0 \leq \beta \leq 1$. Note that Lemma 1.1 of Berry and Fristedt (1979) guarantees the existence of an optimal strategy.

THEOREM 2.2. *For all $\beta \in [0, 1]$, if the discount sequence A is regular, then, for all θ_1 and π, there exists an optimal strategy under which arm 1 is never pulled in the first stage.*

PROOF. Let $\mu = E(\theta_2|\pi)$. We give the proof for the case $\theta_1 > \mu$. The case $\theta_1 \leq \mu$ proceeds in the same manner.

Let t_i equal 1 or 2 according as arm 1 or arm 2 is used on the ith pull when following a given strategy. Let Ω_n denote the set of all regular discount sequences $(\alpha_1, \alpha_2, \ldots)$ satisfying the condition $\gamma_{n+1} = 0$. The proof is by induction on n. Clearly, the result holds for every member of Ω_1. Assume it holds for every member of Ω_{n-1}. Consider $A \in \Omega_n$. If it is optimal to set $t_1 = 2$, then, by the induction hypothesis, it is optimal to not use arm 1 in the first stage. Suppose then that $t_1 = 1$ is uniquely optimal. Then, by the induction hypothesis, $t_2 = 1$ is impossible when $K \geq 2$. Furthermore, if $t_1 = t_2 = 1$, then setting $K = 0$ dominates setting $K = 1$, so it must be optimal to take no pulls in the first stage and to choose arm 1 for the second stage.

The only case remaining is that for some optimal strategy τ^* there is a random $L \geq 1$ such that the first stage consists of $L + 1$ observations, with $t_1 = 1$ and $t_2 = \cdots = t_{L+1} = 2$. We will now show that τ^* is not better than the following modification τ of τ^*: Set $t_1 = \cdots = t_L = 2$ and make the $(L + 1)$st pull in the second stage. To do this, we compare the expected utilities of τ and τ^*,

$$u(\tau^*) = E\left\{\beta\theta_1\alpha_1 + \beta\sum_{i=1}^{L} Y_i\alpha_{i+1} + H(L)\gamma_{L+2}\right\}$$

and

$$u(\tau) = E\left\{\beta\sum_{i=1}^{L} Y_i\alpha_i + H(L)\gamma_{L+1}\right\},$$

where $H(m) = E[\max\{\theta_1, E(\theta_2|Y_1, \ldots, Y_m)\}]$ and, as before, Y_i denotes the ith observation on arm 2.

Let $D \equiv u(\tau) - u(\tau^*)$. Using the fact that $\alpha_1 = 1$, some algebra will verify that

$$D = E\left\{\sum_{i=1}^{L}(\beta Y_i - \theta_1)(\alpha_i - \alpha_{i+1}) + [H(L) - \theta_1]\sum_{i=L+1}^{\infty}(\alpha_i - \alpha_{i+1})\right\} + (1 - \beta)\theta_1.$$

Define $Z_i = (\beta Y_i - \theta_1)I_{\{i \leq L\}} + [H(L) - \theta_1]I_{\{i > L\}}$ and let $B_i = E(Z_i)$. Then

$$D = E\left\{\sum_{i=1}^{\infty} Z_i(\alpha_i - \alpha_{i+1}) + (1 - \beta)\theta_1\right\}$$

$$= B_1\alpha_1 + \sum_{i=1}^{\infty}(B_{i+1} - B_i)\alpha_{i+1} + (1 - \beta)\theta_1.$$

If we define $g_1 = u(\tau^*) - \beta\theta_1\alpha_1 - \theta_1\gamma_2$, then g_1 is $u(\tau^*)$ minus the expected utility for the strategy that takes exactly one first-stage pull on arm 1 and then chooses arm 1 for the second stage. Since τ^* is optimal, $g_1 \geq 0$. However,

$$g_1 = E\left\{\sum_{i=1}^{\infty} Z_i\alpha_{i+1}\right\} = B_1\gamma_2 + \sum_{i=1}^{\infty}(B_{i+1} - B_i)\gamma_{i+2}.$$

Now, if $\alpha_{i+2} > 0$ and if A is regular, then it is easy to show that $(\alpha_1/\gamma_2) \leq (\alpha_{i+1}/\gamma_{i+2})$. Hence,

$$0 \leq (\alpha_1/\gamma_2)g_1 \leq B_1\alpha_1 + \sum_{i=1}^{\infty}(B_{i+1} - B_i)\alpha_{i+1} = D - (1-\beta)\theta_1 \leq D.$$

The second inequality follows from the fact that $B_{i+1} \geq B_i$. To see this, consider

$$B_{i+1} - B_i = E(Z_{i+1} - Z_i) = E\left[I_{[L=i]}\{H(L) - \beta Y_i\} + I_{[L>i]}\beta\{Y_{i+1} - Y_i\}\right].$$

Note that the expected value of the last term is 0. As for the first term, conditioning on $L = i$, we have

$$\begin{aligned} E\left[I_{[L=i]}\{H(L) - \beta Y_i\}|L = i\right] &\geq E\{\max\{\theta_1, E(\theta_2|Y_1, \ldots, Y_i)\} - Y_i|L = i\} \\ &= E\{\max\{\theta_1, E(\theta_2|Y_1, \ldots, Y_i)\} - \theta_2|L = i\}, \end{aligned}$$

which, by Jensen's inequality, is greater than or equal to

$$\max\{E(\theta_1|L = i), E\left[E(\theta_2|Y_1, \ldots, Y_i)|L = i\right]\} - E(\theta_2|L = i). \tag{2.4}$$

But, $E[E(\theta_2|Y_1, \ldots, Y_i)|L = i] = E(\theta_2|L = i)$, since the event $L = i$ is an event in the σ-field generated by $Y_1, \ldots, Y_i$. Thus, expression (2.4) is not less than 0, and, therfore, $B_{i+1} \geq B_i$.

It follows that $u(\tau) - u(\tau^*) \geq 0$. Hence, for $A \in \Omega_n$, there exists an optimal strategy under which arm 1 is never pulled in the first stage. This completes the induction. The result now follows from a suitable modification of Lemma 1.1 of Berry and Fristedt (1979). □

Note that the effective discount sequence A_β will not, in general, be regular, even if A is. Theorem 2.2 states that even in such cases, for all $\theta_1 \in [0,1]$ and π, there exists an optimal strategy under which arm 1 is never pulled in the first stage as long as the associated distribution discount sequence is regular.

On the surface, it may seem as though Theorem 2.2 contradicts Theorem 2.1. However, Theorem 2.1 applies to fixed discount sequences, insofar as A is fixed. Theorem 2.2 applies to discount sequences that are random, in that A_β depends on the (data dependent) random quantity K. The choice of K through an optimal strategy is such that the result of Theorem 2.2 does indeed hold. [A discussion of random discount sequences of a different sort appears in Berry (1983).]

As mentioned, Theorem 2.1 says that when $\beta = 1$, then regularity of A is both necessary and sufficient for an optimal strategy to exist that never pulls the first arm in the first stage. When $\beta < 1$, regularity is not necessary: For example, if $\beta = 0$, then for all θ_1, π and any distribution Q on N, an optimal strategy exists that never pulls arm 1 in the first stage. A somewhat more elaborate example is as follows.

EXAMPLE 2.2. Suppose Q is such that $P(N = 3) = b = 1 - P(N = 1)$. This yields a distribution discount sequence of the form $A = (1, b, b, 0, 0, \ldots)$, which is not regular when $0 < b < \frac{1}{2}$. However, for any given θ_1 and π, if $\beta < bE[\max\{\theta_1, E(\theta_2|Y_1)\}]/[\theta_1 + E(\theta_2)(1-b)]$, then there is an optimal strategy

under which arm 1 is never pulled in the first stage. This may be shown by evaluating the utility of all possible strategies and eliminating those that are suboptimal.

3. Additional properties of optimal strategies. In this section we discuss further the properties of optimal strategies of the two-stage bandit when A is regular, expecially in relation to similar properties of the "standard" bandit of Berry and Fristedt (1979). We focus on the optimal first-stage length, K^*. A more detailed discussion of these properties can be found in Clayton and Witmer (1987).

Consider first the standard bandit, or equivalently, the two-stage bandit with $\beta = 1$. If there exists an n' such that $\gamma_{n'} = 0$, then the problem is "truncated" in the sense that $P(K^* \leq n') = 1$. If, however, $\beta = 1$ and $\alpha_n > 0$ for all n, then there exist priors π such that $P(K^* > n) > 0$ for all n.

We seek those situations for which truncation of K^* occurs in the two-stage bandit when $\beta < 1$. Clearly, this will occur if there exists an n' such that $\gamma_{n'} = 0$. Similarly, if $P(\theta_2 \leq \theta_1) = 0$ or $P(\theta_2 \geq \theta_1) = 0$, then K^* is truncated. The following theorem covers less trivial cases.

THEOREM 3.1. *If $\beta < 1$ and if A is regular such that $\gamma_n > 0$ for all n, then there exists an $n' < \infty$ such that $P(K^* < n') = 1$.*

PROOF. If $\theta_1 = 0$, then $K^* = 0$, and the assertion holds. Suppose that $\theta_1 > 0$ and that observations $Y_1, \ldots, Y_n$ have been made on arm 2. If the first stage is ended immediately, then the utility of doing so is

$$\sum_{i=1}^{n} \beta Y_i \alpha_i + \gamma_{n+1} \max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_n)\}. \tag{3.1}$$

If, instead, one additional observation is made on arm 2, and then the first stage is ended, then the expected utility is, conditional on $Y_1, \ldots, Y_n$,

$$\begin{aligned} &\sum_{i=1}^{n} \beta Y_i \alpha_i + \beta \alpha_{n+1} E(Y_{n+1} | Y_1, \ldots, Y_n) \\ &\quad + \gamma_{n+2} E\{\max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_{n+1})\} | Y_1, \ldots, Y_n\}. \end{aligned} \tag{3.2}$$

Let $\Delta_n(Y_1, \ldots, Y_n)$ denote the quantity in (3.2) minus that in (3.1). Theorem 7.2.5 of Ferguson (1967) may be modified to show that if there exists some n' such that for all $n \geq n'$, $\Delta_n(Y_1, \ldots, Y_n) < 0$ almost surely, then $P(K^* \leq n') = 1$. We show that such an n' exists.

Note that $\gamma_n > 0$ for all n implies $\alpha_n > 0$ for all n. We may therefore write

$$\begin{aligned} \Delta_n(Y_n)/\alpha_{n+1} &= \beta E(\theta_2 | Y_1, \ldots, Y_n) - \max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_n)\} \\ &\quad + (\gamma_{n+2}/\alpha_{n+1})\big(E\big[\max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_{n+1})\} | Y_1, \ldots, Y_n\big] \\ &\qquad - \max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_n)\}\big) \\ &\leq (\gamma_2/\alpha_1)\big(E\big[\max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_{n+1})\} | Y_1, \ldots, Y_n\big] \\ &\qquad - \max\{\theta_1, E(\theta_2 | Y_1, \ldots, Y_n)\}\big) - (1 - \beta)\theta_1. \end{aligned} \tag{3.3}$$

The preceding inequality results from the fact that, since A is regular, $\gamma_{n+2}/\alpha_{n+1} \leq \gamma_2/\alpha_1$, and from the fact that

$$\begin{aligned}
&\beta E(\theta_2|Y_1,\ldots,Y_n) - \max\{\theta_1, E(\theta_2|Y_1,\ldots,Y_n)\} \\
&\quad = \beta\left[E(\theta_2|Y_1,\ldots,Y_n) - \max\{\theta_1, E(\theta_2|Y_1,\ldots,Y_n)\}\right] \\
&\qquad -(1-\beta)\max\{\theta_1, E(\theta_2|Y_1,\ldots,Y_n)\} \\
&\quad \leq -(1-\beta)\theta_1.
\end{aligned}$$

It will suffice to show that there exists an n' such that $n \geq n'$ implies that the right-hand side of inequality (3.3) is less than 0 almost surely. But this is equivalent to the "binomial sequential test" of Ray (1965), Example 5.2, and following the discussion preceding equation 5.15 of Ray (1965), it suffices to show that $\lim_{n\to\infty}\operatorname{ess\,sup}\operatorname{Var}(\theta_2|Y_1,\ldots,Y_n) = 0$. This follows from a proof similar to that of Theorem 4.1.4 of Chung (1974). □

Note that n' may be a poor bound on K^*. If A is of a special form (say, geometric), then an approach like that in Ray (1965), Example 5.3, could be used to provide a better bound.

The previous theorem and the preceding discussion may be put informally as follows: Except in degenerate situations, K^* is truncated if and only if $\beta < 1$. Using this fact and the technique of backward induction [DeGroot (1970), page 277], we can prove Theorems 3.2 and 3.3. Theorem 3.2 extends Theorem 3.1 of Berry and Fristedt (1979) and establishes a generalized form of monotonicity for the two-stage bandit, whereas Theorem 3.3 confirms the intuition that we tend to take more first-stage observations when β is large than when β is small.

THEOREM 3.2. *If π' is strongly to the right of π [as defined in Berry and Fristedt (1979), Definition 3.2], and if A is regular, then $u(\pi', \beta, \theta_1) \geq u(\pi, \beta, \theta_1)$, where $u(\pi, \beta, \theta_1)$ denotes the maximal utility for given π, β and θ_1. Also, $u(\pi, \beta, \theta_1)$ is an increasing function of β and θ_1.*

THEOREM 3.3. *If $\beta < \beta' \leq 1$, then $K^*(\beta) \leq K^*(\beta')$.*

4. Discussion. The ideas presented here can be applied in the setting of a clinical trial in which two treatments, one standard and one new, are to be compared. The discount factor β quantifies the relative importance of effectively treating patients in the trial to effectively treating patients after the trial. This idea is more fully considered in Clayton and Witmer (1986).

In a related work, Chernoff and Petkau (1985) have considered the "ethical costs" associated with assigning the apparently inferior treatment to a patient when the treatment responses are normally distributed. In the setting in which treatment yields dichotomous response and both θ_1 and θ_2 are unknown, Simons (1986) has parameterized such ethical costs. These costs have a similar effect to that of our β.

When N is unknown, one is tempted to avoid the use of Q by replacing N in (2.1) with $E(N|Q)$ and proceeding as if N were known. In Clayton and Witmer

(1986), we investigate the effect of such practice in the nonsequential setting (i.e., when K is chosen before the first pull). We show that, if Q is regular, then such a method yields a strategy that is nearly optimal. We conjecture that this is true in the sequential setting as well.

Acknowledgment. The authors thank the referee for helping to clarify the exposition.

REFERENCES

BERRY, D. A. (1983). Bandit problems with random discounting. In *Mathematical Learning Models —Theory and Algorithms* (H. Herkenrath, D. Kalin and W. Vogel, eds.) 12–25. Springer, New York.

BERRY, D. A. and FRISTEDT, B. (1979). Bernoulli one-armed bandits—arbitrary discount sequences. *Ann. Statist.* **7** 1086–1105.

BERRY, D. A. and FRISTEDT, B. (1985). *Bandit Problems: Sequential Allocation of Experiments.* Chapman and Hall, London.

CANNER, P. L. (1970). Selecting one of two treatments when the responses are dichotomous. *J. Amer. Statist. Assoc.* **65** 293–306.

CHERNOFF, H. and PETKAU, A. J. (1985). Sequential medical trials with ethical cost. In *Proc. of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **2** 521–537. Wadsworth, Monterey, Calif.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic, New York.

CLAYTON, M. K. and WITMER, J. A. (1986). Generalized two-stage decision problems. Technical Report 774, Dept. Statistics, Univ. Wisconsin.

CLAYTON, M. K. and WITMER, J. A. (1987). Some theorems for two-stage bandits. Technical Report 803, Dept. Statistics, Univ. Wisconsin.

DEGROOT, M. H. (1970). *Optimal Statistical Decisions.* McGraw-Hill, New York.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision–Theoretic Approach.* Academic, New York.

PETKAU, A. J. (1978). Sequential medical trials for comparing an experimental with a standard treatment. *J. Amer. Statist. Assoc.* **73** 328–338.

RAY, S. N. (1965). Bounds on the maximal sample size of a Bayes sequential procedure. *Ann. Math. Statist.* **36** 859–878.

SIMONS, G. (1986). Bayes rules for a clinical-trials model with dichotomous responses. *Ann. Statist.* **14** 954–970.

WITMER, J. A. (1986). Bayesian multistage decision problems. *Ann. Statist.* **14** 283–297.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706

DEPARTMENT OF MATHEMATICS
OBERLIN COLLEGE
OBERLIN, OHIO 44074-1094