Using Patient Characteristics and Attitudinal Data to Identify Depression Treatment Preference Groups: A Latent-Class Model

Jennifer A. Thacher *
Department of Economics, University of New Mexico

Edward Morey †
Department of Economics, University of Colorado

W. Edward Craighead
Department of Psychology, University of Colorado

January 28, 2004

*Correspondence to Jennifer Thacher, Department of Economics, University of New Mexico, MSC05 – 3060, Albuquerque, NM 87108. Email: jthacher@unm.edu. Phone: 505-277-1965. Fax: 505-277-9445.
†Thacher and Morey are equal authors and rotate primary author position.
Abstract

The model identifies and characterizes groups of patients who share similar attitudes towards treating depression. The results predict the probability of preference group membership on the basis of observable characteristics and answers to attitudinal questions. Understanding the types of preference groups that exist and a patient’s probability of membership in each of the groups can help clinicians tailor the treatment to the patient and may increase patient adherence. 104 depressed patients completed a survey on attitudes towards treatment of Major Depressive Disorder. Analysis showed that treatment preferences vary among depressed patients. Two classes were identified that differ in their sensitivity to treatment costs and side effects. One class cared primarily about treatment effectiveness; side effects and the cost of treatment had little impact on this class’s treatment decisions. Another class was highly sensitive to cost and side effects. Younger, male, or less educated individuals were more likely to be sensitive to treatment costs and side effects.

Short title: Treatment Preference Groups

Keywords: Major Depressive Disorder, Patient Care, Patient Acceptance of Health Care
Introduction

Understanding how preferences vary is important to behavioral researchers, particularly those studying health-care. Several studies suggest that treatment preferences can differ significantly between patients (Brundage et al. 2001; Nease et al. 1995; Tsevat et al. 1998). Patient preferences can affect treatment decisions and outcomes (Wu et al. 2001). Montgomery et al. (2001) found that treatment recommendations vary when patient preferences are taken into account. A number of studies in mental health have examined patient preferences. Revicki and Wood (1998), Dwight-Johnson et al. (2000), O’Brien et al. (1995), Cooper et al. (2000), and Sestoft et al. (1998) examined patient preferences over depression treatment programs. Kremer and Gesten (2003) examined preferences for managed-care psychotherapy. Walburn et al. (2001) studied preferences for antipsychotic medications.

It has been observed that a significant share of individuals with Major Depressive Disorder (MDD) discontinue treatment prematurely (Demyttenaere 1997; McCombs et al. 1990; Simon et al. 1996; Thompson et al. 1996). For example, Lin et al. (1995) found that 28% of primary care patients stop taking anti-depressants within one month of beginning treatment; 44% stop within three months. Myers and Branthwaite (1992) found a compliance rate of 68% after three weeks in depressive patients; this rate declines to 50% after 12 weeks. A mismatch between the treatment method and patient treatment preferences may explain this observed behavior. This observation provides a rationale for learning more about depression treatment preferences.

The standard gamble, time-trade-offs, choice questions, and willingness-to-pay are currently the primary methods of examining patient preferences (Green et al. 2000). These methods are discussed below. This paper develops an additional method for learning about treatment preferences: a latent-class attitudinal model. We apply the model to depression treatment preferences.

The latent-class model presented here is based on answers to a set of Likert Scale attitudinal questions and examines mental health treatment preferences. Consider the following attitudinal question from our application:

How important to you is the use of anti-depressants in choosing a depression treatment program? (Not Important at All, Not Very Important, Somewhat Important, Pretty Important, Very Important)
A statistical analysis of individual responses to a set of these type of questions allows the researcher to identify and characterize various treatment preference classes. For example, in our application it is used to identify groups that differ in their sensitivity to various treatment attributes, such as type of treatment, cost, effectiveness, and side-effects (weight gain, reduced sex drive, and inability to orgasm). To which treatment preference class an individual belongs is latent/unobserved. The model estimates the probability of being in distinct preference classes as a function of observed characteristics of the individual. In the application, age, education, and gender are examined. Treatment preferences are assumed homogenous within each latent-class, but vary significantly across classes with respect to cost, side-effects, and type of treatment.

Standard references to latent-class models include Bartholomew and Knott (1999), Titterington et al. (1985), and Wedel and Kamakura (2000). Our model is similar to Bandeen-Roche et al. (1997) in its use of observable individual characteristics. Their model is not estimated with attitudinal data. A few researchers have applied latent-class models to attitudinal data (Clogg and Goodman 1984; De Menezes and Bartholomew 1996; Eid et al. 2003; McCutcheon 1987; McCutcheon and Nawojczyn 1995; Yamaguchi 2000). However, to our knowledge, there are no applications of this method to patient preferences. Researchers that have previously analyzed mental-health treatment preferences on the basis of attitudinal data, compared the attitudinal answers of different demographic groups; they did not use statistical models (Cooper et al. 2000; Sestoft et al. 1998).

In comparison, the time-trade-off method asks how many life-years an individual will give up in order to experience perfect health rather than depression (Wells and Sherbourne 1999). Choice questions present a patient with hypothetical choices between treatment methods. Each treatment can vary in effectiveness, cost, and side effects (O’Brien et al. 1995; Johnson et al. 2000; Ryan et al. 2000). Willingness-to-pay methods ask the patient whether she would pay a certain amount to eliminate her depression or prevent certain side effects. O’Brien et al. (1995) is an application of the willingness-to-pay method to depression. The standard gamble method asks an individual to choose between depression with certainty, and a gamble between no depression and death. In the gamble, depression and death each occur with a certain probability. An individual’s gamble score is the probability of death that makes her indifferent between depression and the gamble. Revicki and Wood (1998) and Wells and Sherbourne (1999) applied this method to mental health.
Different methods exist to treat depression. Understanding the types of preference classes that exist and a patient’s probability of membership in each of the classes can help clinicians better tailor the treatment to the patient, potentially improving patient adherence. Our results can be used in a clinical setting to determine a patient’s probability of belonging to a treatment preference class solely on the basis of her observed socioeconomic characteristics. If the resources are available to give the patient a short attitudinal survey (e.g., on a computer notepad in the waiting room), patients can be allocated to a preference class with an even higher degree of probability.

Methods

This paper applies the latent-class model to analyze attitudes toward different elements of depression treatment. Up to three preference classes are allowed.

Sample

The data come from a survey of depressed adults seeking treatment for a new episode of MDD at a HMO mental health clinic in Colorado. The sample includes both individuals seeking treatment for the first time and individuals previously treated for MDD. Financially independent individuals age 18 and older and diagnosed by clinicians as suffering from MDD were eligible to participate. The survey asks patients about their preferences over the elements of depression treatment programs, including treatment effectiveness, use of anti-depressants, number of hours of psychotherapy per month, out-of-pocket cost, and the presence of three possible side effects (weight gain, reduced sex drive, and inability to orgasm). Patients report on a scale of 1 to 5 the importance of each of these treatment elements in their treatment choice.¹

One hundred seven patients filled out a survey that included ten attitudinal questions.² All cases with missing data were deleted, leaving a sample size of 104. Women constituted 75% of the sample.

¹Possible answers range from “Not Important at All” (1) to “Very Important(5)”. In order to limit the number of parameters estimated, these five possible answers were collapsed into three variables. This is a common procedure. For example, see De Menezes and Bartholomew (1996) and Yamaguchi (2000). Results were robust to a number of different ways of collapsing the data.

²Only data from seven of the attitudinal questions were used as several of the attitudinal questions were similar.
the sample. Eighty-one percent of the sample was White, Non-Hispanic. The average age was 40 (s.d.=11); the youngest was 18 and the oldest 74. The majority identified some college as their highest completed level of education. The average household income, based on the midpoint of income ranges, was $53,738 (s.d.=30,516). Forty-five percent of the sample were receiving their first treatment for depression.

The Latent-Class Model

The model assumes that the population consists of a number of different preference classes. An individual’s preference class is unobserved/latent. The researcher observes an individual’s set of answers to the attitudinal questions (the individual’s response pattern) and characteristics of the individual such as gender and retirement status. The response patterns of individuals from the same preference class are more correlated with each other than with individuals in other classes; individuals of the same type answer similarly. Latent-class models assume that once you have controlled for class membership, attitudinal responses are independent.

The estimation goal is to find the response probabilities (the probability that an individual in a certain preference class gives a particular answer to an attitudinal question) and unconditional class probabilities (the probability that an individual belongs to a particular preference class given her observable characteristics) that are most likely, given the response patterns. For example, the unconditional probability might represent the probability that an older woman belongs to a particular treatment preference class; this probability does not depend on her specific answers to the attitudinal questions. All individuals with the same observed characteristics possess the same unconditional probability of belonging to a particular preference class. The unconditional probabilities are not known a priori. The conditional membership probabilities can also be calculated; these are the probability that an individual belongs to a particular class given her observable characteristics and specific answers to the attitudinal questions.

The ln likelihood function for a $C$-class model for the data in this sample was:

$$\ln L = \sum_{i=1}^{104} \ln \left[ \sum_{c=1}^{C} \Pr(c : z_i) \prod_{q=1}^{7} \prod_{s=1}^{3} \left( \pi_{qs|c} \right)^{x_{qs|i}} \right],$$  \hspace{1cm} (1)

where $\pi_{qs|c}$ is the conditional probability that individual $i$ answers $s$ on question $q$, $\Pr(c : z_i)$ is

Thacher 6 Treatment Preference Groups
the unconditional probability that an individual with characteristics \( z_i \) belongs to class \( c \), and \( x_{iqs} \) is a dummy variable that shows whether individual \( i \) chose \( s \) on question \( q \).

Estimation is with the E-M algorithm (Dempster et al. 1977; Bartholomew and Knott 1999). The results for this paper were programmed in Gauss.

Factor analysis and cluster analysis are two other possible methods for characterizing individuals on the basis of their answers to a set of attitudinal questions. Factor analysis is of limited used in this application because it does not identify discrete preference groups; rather, each individual has an unique factor variable score. Since the purpose of this study is to capture trends in treatment preferences and to identify which issues are most important, this approach is of limited value. Cluster analysis, like latent class analysis, allocates respondents into a finite number of groups on the basis of their answers to the set of questions. However, unlike the latent class approach, group membership is assumed deterministic; individuals are allocated to groups as if the researcher knows with 100% certainty to which group individuals should belong. This is a strong assumption. Latent class analysis identifies distinct groups but assigns probabilities to belonging to each group, thus highlighting individual differences.

Data Analysis

A number of criteria were used to assess the degree of fit of the model, including the estimated conditional probabilities, the entropy statistic, and the \( AIC \), \( CAIC \), and \( AIC_C \) information criteria. If the number of classes fit the data well, the estimated conditional membership probabilities should be close to one or zero; individuals should be assigned to one of the classes with high probability. The entropy statistic measures the degree to which the model successfully separates individuals into classes. This measure is bounded by 0 and 1. A value close to 1 indicates that the model successfully separates individuals into classes (Wedel and Kamakura 2000). Following previous research, the \( AIC \), \( CAIC \), and \( AIC_C \) information criteria are also used as indicators for selecting the correct number of preference classes (Akaike 1974; Bozdogan 1987; Hurvich and Tsai 1989). These information criteria are essentially log-likelihood scores with a correction factor for
sample size and number of parameters. The best fitting model minimizes the information criteria.\textsuperscript{3} It was expected that the best model would have high estimated conditional probabilities, a high entropy score, and low information scores relative to other models.

**Results**

Table 1 shows the ln likelihood value, information criteria, and entropy scores for the one-, two-, and three-class models. All models were run with 200 random starts. By all applicable measures, the one-class (no heterogeneity) model provides the worst fit. In the case of depression treatment it is inappropriate to assume strict preference homogeneity; individuals have different preferences for treating their depression. Based on the $AIC_C$ and $CAIC$ criteria, the two-class model performs the best, suggesting the presence of two depression treatment classes. Based on the $AIC$ information criteria, the three-class model performs the best, suggesting the presence of three depression treatment classes. Both models score high on the entropy measure, indicating that it successfully separates individuals into classes. The $CAIC$ and $AIC_C$ are more conservative measures than the $AIC$ with a larger penalty for increases in the number of parameters.

The interpretation of the two- and three-class models are fairly similar. In both the two- and three-class models, all preference groups consider the effectiveness of treatment most important in choosing a treatment plan. Groups differ in how they feel about other aspects of treatment, such as side-effects and cost. In the two-class model, one group cares primarily about the effectiveness of treatment; other components of treatment such as side-effects and cost are of relatively little importance. Call this group the *Not Sensitive* group; this group will likely participate in treatment that is expected to be effective, regardless of other costs. The other group, while it still considers effectiveness most important, also considers the side-effects to be very important. Call this group the *Sensitive* group. Individuals in this group consider other aspects of treatment, such as sexual

\textsuperscript{3}Definitively determining the correct number of preference classes requires a statistical test of whether one model provides a better fit than a competing model with a different number of classes. Unfortunately, no statistical test exists; one can examine test statistics, but their distributions are unknown. For example, one can calculate the likelihood ratio statistic for $C$ versus $C + 1$ classes but the regularity conditions used to prove that this statistic has a $\chi^2$ distribution will be violated.
side effects to be very important. The presence of these costs are more likely to hinder their participation in treatment.

If a three-group model is allowed, the results are similar. Again one sees the presence of a Sensitive and Not Sensitive group. In addition, one finds a third group, Somewhat Sensitive. Members of this group again consider effectiveness to be most important. Other aspects of treatment are less important to them than to the Sensitive group but more important than that of the Not Sensitive group. In addition, the weight gain side effect is relatively more important to this group while the sex drive side effect is relatively less important.

On the basis of the information criteria, we have chosen to report results for the two-class model and to show the effects of adding covariates for this model. However, we again emphasize the similar results between the two- and three-class models. Not surprisingly, adding another class allows for more detailed differences in treatment preferences to arise. We suspect that with a larger sample size, a model with greater preference differences would be chosen.

Based on a likelihood ratio test ($\chi^2 = 18.61$), a model that included information on an individual’s age, gender, and education level statistically dominated a two-class model that did not include all this information. This is also true for all sub-models.

Table 2 reports some example estimated response probabilities; these help illustrate how the two classes differ. Consider the model’s predictions of how individuals in the two classes will answer the question “How important is treatment effectiveness?” Though all three classes consider treatment effectiveness important, an individual in the Not Sensitive class has a 93% probability of answering that treatment effectiveness is Very Important. In contrast, individuals in the Sensitive class are predicted to give the same answer with 70% probability.

The model predicts with an 53% probability that an individual in the Sensitive class will characterize a reduced sex drive as Very Important; individuals in the Not Sensitive class have an extremely small probability of giving the same answer. In fact, with 80% probability, individuals in the Not Sensitive class will describe the reduced sex-drive side effect as either Not at All or Not Very Important in making treatment decisions. Similarly, individuals in the Sensitive class have a high probability (56%) of saying that the weight gain side effect is Very Important. This probability is much lower for someone in the Not Sensitive class. A similar story holds for cost,
although individuals in this sample are less sensitive to cost than medication side effects.

Table 3 reports the unconditional membership probabilities for individuals who differ by gender, age category, and education level. These numbers are the predicted probabilities of class membership if you only observe characteristics of the individual but have no information about an individual’s answers to the attitudinal questions. For example, the model predicts that a young woman without a college degree is more likely to be very sensitive to treatment side effects and costs than an older woman with a college degree (83% versus 37%). Thus, without any more information about a woman than her age and education level, the most reasonable assumption to make about an older woman with a college degree who comes into a mental health clinic for treatment for depression is that she primarily cares about treatment effectiveness.

The model also predicts that regardless of gender, younger individuals care more about treatment side effects than do older individuals. Regardless of age or education level, men are more likely than women to be very sensitive to treatment costs and side effects. Young males without a college degree have the highest probability of caring about costs and side effects (95%). Younger individuals without college degrees are more likely to be sensitive to costs and side-effects than their counterparts with college degrees. Although education decreases treatment barriers, however, it fails to eliminate them.

Having a new patient answer the set of seven attitudinal questions dramatically increases the probability of identifying to which preference class the individual belongs. The conditional membership probabilities are 95% or higher for 90% of the sample. For illustrative purposes, consider the case of older women with college degrees. These individuals have high unconditional probabilities of belonging to the Not Sensitive group but when their specific answers to the attitudinal questions are taken into account, some of these same individuals are more likely to belong to another group. The conditional probabilities that these women belong to the Not Sensitive class range from 0 to 100%. For example, one woman in this class answered that all the side effects and effectiveness were Very Important while cost, use of anti-depressants, and number of hours of therapy were Somewhat Important. Based on these answers, this woman had a conditional probability of 100% of being in the Sensitive class and a zero percent probability of being in the Not Sensitive group.

\(^4\) We did not allow for more than three personal characteristics because of concerns about the number of people in each group. Other age and education cutoff points were investigated but were not significant.

Thacher 10 Treatment Preference Groups
class. Thus, while the unconditional probabilities can be good indicators on average of a person’s preference class, the incorporation of attitudinal data allows for much more accurate prediction.

Discussion

This paper presented a latent-class model estimated with attitudinal data that can be used to improve researchers’ and clinicians’ understanding of patient preferences. In contrast to other types of techniques such as standard gamble, time-trade-off, willingness-to-pay, and choice questions, the presented model utilizes Likert Scale attitudinal data, which is relatively quick and easy data to obtain.

The application showed that treatment preferences vary among depressed patients. Two classes were identified that differ in their sensitivity to treatment costs and side effects. One class cared primarily about treatment effectiveness; side effects and the cost of treatment had little impact on this class’s treatment decisions. Another class was highly sensitive to cost and side effects; each of these variables had a large impact on this class’s treatment decisions, possibly affecting both the initial treatment decision and adherence to the treatment program. Adding observable individual characteristics produced a better fitting model. For example, individuals who are younger, male, or less educated were more likely to be sensitive to treatment costs and side effects.

By applying the latent class model, investigators can identify distinct classes that differ in their treatment preferences and concerns. They can generate estimates of the size of these classes for different patient populations.

We believe that this method has a number of possible applications. First, the unconditional probabilities estimated in this paper can provide guidance for clinicians about how treatment preferences vary among demographic groups. Second, by having patients answer a brief set of attitudinal questions, conditional probabilities could be quickly calculated, allowing identification of a patient’s specific treatment concerns. As was shown in the application, knowledge of the conditional probabilities allow the clinician to assign a patient to a treatment preference group with a high degree of confidence. Knowing to which preference group a patient belongs is useful information to the clinician. It allows the clinician to focus limited appointment time on addressing the patient’s concerns. In addition, having more information about patient preferences should allow

Thacher 11 Treatment Preference Groups
the clinician to be of greater assistance in helping the patient to pick the best treatment plan. It is our expectation that treatment plans that better reflect patient preferences will result in increased adherence to the treatment plan. Testing this hypothesis is an area of future research.
Acknowledgements. We thank Kaiser Permanente for allowing us to conduct the study and Arne Beck, Cathy Bartsch, and Carolee Nimmer for sponsoring the project. Special thanks to the patients who participated in the study and the clinicians who recruited them, in particular Bonnie Mucklow, Stew Nyholm, John Romeo, Nancy Rogers, and Marsha Visscher. Thanks to Bill Breffe, Roger von Haefen, Nick Flores, Luke Rodgers, and Don Waldman for valuable comments and suggestions.
References


Thacher 16 Treatment Preference Groups
Wells K, Sherbourne C. 1999. Functioning and utility for current health of patients with depression or chronic medical conditions in managed primary care practices. Arch Gen Psychiatry 56:897–904.


Table 1: Goodness of Fit Measures Suggest Heterogeneous Treatment Preferences

<table>
<thead>
<tr>
<th>Model</th>
<th>lnL</th>
<th>AIC</th>
<th>CAIC</th>
<th>AICc</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 group</td>
<td>−698.24</td>
<td>1424.49</td>
<td>1461.65</td>
<td>1429.94</td>
<td>n/a</td>
</tr>
<tr>
<td>2 groups</td>
<td>−649.21</td>
<td>1356.41</td>
<td>1433.37</td>
<td>1381.89</td>
<td>0.9524</td>
</tr>
<tr>
<td>3 groups</td>
<td>−617.33</td>
<td>1322.65</td>
<td>1439.43</td>
<td>1394.03</td>
<td>0.9233</td>
</tr>
</tbody>
</table>

Table 2: Groups Differ by Sensitivity to Treatment

<table>
<thead>
<tr>
<th>Predicted Response Probabilities</th>
<th>Not Sensitive</th>
<th>Sensitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \pi_{qs</td>
<td>1} )</td>
</tr>
<tr>
<td><strong>Sex-Drive Side Effect</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at All/Not Very Important</td>
<td>80%</td>
<td>0%</td>
</tr>
<tr>
<td>Somewhat/Pretty Important</td>
<td>16%</td>
<td>47%</td>
</tr>
<tr>
<td>Very Important</td>
<td>4%</td>
<td>53%</td>
</tr>
<tr>
<td><strong>Weight-Gain Side Effect</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at All/Not Very Important</td>
<td>23%</td>
<td>11%</td>
</tr>
<tr>
<td>Somewhat/Pretty Important</td>
<td>54%</td>
<td>33%</td>
</tr>
<tr>
<td>Very Important</td>
<td>23%</td>
<td>56%</td>
</tr>
<tr>
<td><strong>Effectiveness</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at All/Not Very Important</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>Somewhat/Pretty Important</td>
<td>7%</td>
<td>29%</td>
</tr>
<tr>
<td>Very Important</td>
<td>93%</td>
<td>70%</td>
</tr>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at All/Not Very Important</td>
<td>43%</td>
<td>13%</td>
</tr>
<tr>
<td>Somewhat/Pretty Important</td>
<td>40%</td>
<td>56%</td>
</tr>
<tr>
<td>Very Important</td>
<td>17%</td>
<td>31%</td>
</tr>
</tbody>
</table>
Table 3: Young, Male, and Less Educated are More Likely to be Sensitive to Treatment Costs and Side-Effects

<table>
<thead>
<tr>
<th>Unconditional Membership Probabilities</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18-40 &amp; 18-40 &amp; 40+ &amp; 40+ &amp;</td>
<td>18-40 &amp; 18-40 &amp; 40+ &amp; 40+ &amp;</td>
</tr>
<tr>
<td></td>
<td>No College College No College College</td>
<td>No College College No College College</td>
</tr>
<tr>
<td></td>
<td>Degree Degree Degree Degree</td>
<td>Degree Degree Degree Degree</td>
</tr>
<tr>
<td>&quot;Not Sensitive&quot; Group</td>
<td>17% 25% 51% 63%</td>
<td>5% 8% 21% 31%</td>
</tr>
<tr>
<td>&quot;Sensitive&quot; Group</td>
<td>83% 75% 50% 37%</td>
<td>95% 92% 79% 69%</td>
</tr>
</tbody>
</table>