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 23, 2005

*Correspondence to Jennifer Thacher, Department of Economics, University of New Mexico, MSC05 – 3060, Albuquerque, NM 87131. Email: jthacher@unm.edu. Phone: 505-277-1965. Fax: 505-277-9445.

†Thacher and Morey are equal authors and rotate primary author position.

Abstract

A latent-class model is used to identify and characterize 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. One hundred four depressed patients completed a survey on attitudes towards treatment of Major Depressive Disorder. Analysis shows that treatment preferences vary among depressed patients. Three classes are identified that differ in their sensitivity to treatment costs and side effects. One class cares primarily about treatment effectiveness; side effects and the cost of treatment have little impact on this class's treatment decisions. Another class is highly sensitive to cost and side effects. A third class is somewhat sensitive to cost and side effects. Younger and male patients are 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 differ significantly across patients (Nease *et al.* 1995; Tsevat *et al.* 1998; Brundage *et al.* 2001). Patient preferences 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 have examined patient preferences over depression treatment programs (O'Brien *et al.* 1995; Revicki and Wood 1998; Sestoft *et al.* 1998; Dwight-Johnson *et al.* 2000; Cooper *et al.* 2000). Kremer and Gesten (2003) examined preferences for managed-care psychotherapy. Walburn *et al.* (2001) studied preferences for antipsychotic medications.

A significant share of individuals with Major Depressive Disorder (MDD) discontinue treatment prematurely (McCombs *et al.* 1990; Simon *et al.* 1996; Thompson *et al.* 1996; Demyttenaere 1997). 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 and have all been applied to learning about patient preferences for depression treatment (O'Brien *et al.* 1995; Revicki and Wood 1998; Wells and Sherbourne 1999; Green *et al.* 2000). We model patient preferences and their heterogeneity in terms of unobserved or latent preference classes. Estimation is based solely on the answers to a set of Likert Scale attitudinal questions that elicit mental health treatment preferences. Consider the following attitudinal question from our application:

> How important to you is the monthly cost of treatment in choosing a depression treatment program? (*Not Important at All, Not Very Important, Somewhat Important, Pretty Important, Very Important*)

We identify three classes of patients 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. Treatment preferences are assumed homogenous within each latent-class, but vary significantly across classes with respect to cost, side effects, and type of treatment. One group cares primarily about completely eliminating their depression; other components of treatment such as side effects and cost were of relatively little importance. This Least Sensitive group will likely participate in treatment that is expected to be fully effective, regardless of other financial and physical costs. A second group (Most Sensitive) considers treatment effectiveness important, but considers sexual side effects to be even more important. The presence of these costs are more likely to hinder their participation in treatment. Members of the third group (Somewhat Sensitive) consider effectiveness to be most important. Side effects are more important to them than the Least Sensitive group but less important than the Most Sensitive group. Age and gender impact group membership.

Standard references to latent-class models include Titterington *et al.* (1985), Bartholomew and Knott (1999), and Wedel and Kamakura (2000). Our model is similar in concept to Bandeen-Roche *et al.* (1997) in its use of observable individual characteristics. Their model was not estimated with attitudinal data. A few researchers have applied latent-class models to attitudinal data (Clogg and Goodman 1984; McCutcheon 1987; McCutcheon and Nawojcyzk 1995; De Menezes and Bartholomew 1996; Yamaguchi 2000; Eid *et al.* 2003; Morey *et al.* 2005). However, to our knowledge, there are no applications of this method to patient preferences. Researchers who 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 (Sestoft *et al.* 1998; Cooper *et al.* 2000).

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 certainty.

Methods

Participants

The data come from a survey of depressed adults seeking treatment for a new episode of Major Depressive Disorder (MDD) at a HMO mental health clinic in Colorado. The sample included 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 asked 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 reported on a scale of one to five 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]

[1] Possible answers ranged 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: Not at All/Not Very Important, Somewhat/Pretty Important, Very Important. This is a common procedure (e.g., De Menezes and Bartholomew (1996) and Yamaguchi (2000)).

[2] Seven of these questions were used in the analysis. Each of the other three questions were very similar

All cases with missing data were deleted, leaving a sample size of 104. Women constituted 75% of 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 (46%) identified some college as their highest completed level of education. The median household income, based on the midpoint of income ranges, was $\$55,000$ while the first and third quartiles respectively were $\$35,000$ and $\$70,000$. Forty-five percent of the sample were receiving their first treatment for depression.

A Latent-Class Model

The model assumes that the population consists of a number of different preference classes. The researcher observes an individual's set of answers to the attitudinal questions (the individual's response pattern) and characteristics of the individual. 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 was to find the most likely 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), given the response pattern. 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 probability can also be derived; this is the probability that an individual be-

to the seven included questions, so added no new information. An example of the survey can be found at www.unm.edu/~jthacher/DepressionSurvey.pdf.

longs 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:

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

where $\pi_{qs|c}$ is the response probability that individual i answers s on question q given that she is in class c, $\Pr(c : \mathbf{z}_i)$ is 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. We imposed the constraint that $\sum_{s=1}^{S} \pi_{qs|c} = 1$.

Estimation was with the E-M algorithm (Dempster *et al.* 1977; Bartholomew and Knott 1999). The results for this paper were programmed in GAUSS (Aptech Systems 1995) and LEM (Vermunt 1997). Because of sample size concerns, we did not run the model for more than three classes. All models were run with 20 random starts.

Results

We find that the best model is one with three treatment preference classes. These preference classes vary in their degree of sensitivity to treatment barriers such as side effects and costs.

Model Selection

In choosing the best model one needs to think about two issues. One is how well the model fits the data. The second issue is choosing the most appropriate number of latent classes.

Both the Pearson and Read-Cressie statistics address the first issue by comparing the expected and actual frequencies of responses (Forman 2003). We report bootstrapped p-values for each of these statistics.[3] The bootstrapped p-value is the proportion of estimated

[3]Since a large number of possible response patterns were not observed, the distributions of the test statistics are not known and had to be bootstrapped (von Daiver 1997; Eid *et al.* 2003). Based on the assumption that the estimated parameters from the latent class model were correct, we generated 600

test statistics larger than the test statistic from the original model. A model that fits the data well should have a bootstrapped p-value larger than 0.05. As is shown in Table 1, based on the bootstrapping technique, the one- and two-class models did not fit the data well while the three-class model did. The small p-value of the Read-Cressie statistic suggests that the one- and two-class models should be rejected. In other words, the bootstrapped Read-Cressie statistic took on a larger value than that of our original statistic less than 0.5% of the time. This suggests that the original statistic is significantly different from its expected value. In contrast, for the three-class case we fail to reject the hypothesis that the original parameter estimates were correct. Therefore, we concluded that the three-class model fits the data well.

Information criteria such as the AIC, $CAIC$, AIC_C, and BIC are often used to determine the appropriate number of classes.(Akaike 1974; Bozdogan 1987; Hurvich and Tsai 1989; Schwarz 1978).[4] These information criteria are essentially log-likelihood scores with a correction factor for sample size and number of parameters. The best fitting model in terms of the number of classes is the one which minimizes the information criteria. Information criteria are often used alone without making use of the Pearson and Read-Cressie statistics. Table 1 reports the ln likelihood value and information criteria for the one-, two-, and three-class models. The three-class model is also considered the best model using the information criteria.

After choosing three classes as the most appropriate model, covariates were incorporated by modeling the unconditional probability as a logit specification:

$$\Pr(c : \mathbf{z}_i) = \frac{e^{\beta_c \mathbf{z}_i}}{\sum_{k=1}^{C} e^{\beta_k \mathbf{z}_i}},\ c = 1, ..., C \tag{2}$$

Based on a likelihood ratio test ($\chi^2 = 25.29$), a model that included information on an

simulated datasets. We re-ran the latent-class model on each simulated dataset and calculated the the Pearson and Read-Cressie statistics for each run.

[4]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 χ^2 distribution will be violated.

individual's age and gender statistically dominated a three-class model that did not include this information.[5]

We interpreted the model using the estimated response probabilities (Table 2) and mean responses across groups (Table 3). We examined both ranking of mean responses within a group and comparison of mean responses across groups. Both tables tell a similar story. All preference groups consider the effectiveness of treatment highly important in choosing a treatment plan. Groups differ in how they feel about other aspects of treatment, such as side effects and cost.

Table 2 reports estimated response probabilities; these help illustrate how the classes differ. Consider the model's predictions of how individuals in each class will answer the question "How important is treatment effectiveness?" Though both classes consider treatment effectiveness important, an individual in the Least Sensitive class has a 93% probability of answering that treatment effectiveness is *Very Important.* In contrast, individuals in the other two classes are predicted to give the same answer with 73% and 68% probability.

The model predicts with a 100% probability that an individual in the Most Sensitive class would characterize a reduced sex drive as *Very Important*; individuals in the Least Sensitive class and Somewhat Sensitive classes have a 4% and 17% change respectively of giving the same answer. Similarly, individuals in the Most Sensitive class have a high probability (77%) of saying that the weight gain side effect is *Very Important.* This probability is much lower for someone in either of the other two classes. A similar story holds for cost, although individuals in this sample appear to be less sensitive to cost than medication side effects.

Suppose individuals are assigned to the class for which they had the highest conditional probability of belonging. This is a reasonable thing to do since the conditional membership probabilities were 95% or higher for 93% of the sample. Table 3 reports the actual means to the attitudinal question for each class. The data in this table tells a similar story to that of the response probabilities both when comparing statistically significant means across classes

[5]Other demographic differences such as education level also seemed important. However, again because of concerns about sample size we limited the number of covariates.

and within a class.

Table 4 reports the unconditional membership probabilities for individuals who differ by gender and age category. These numbers are the predicted probabilities of class membership if you only observe the age and gender of the individual but have no information about an individual's answers to the attitudinal questions. For example, the model predicts that a younger woman is more likely to be very sensitive to treatment side effects and costs than an older woman (35% versus 21%). Thus, without any more information about a woman than her age, the most reasonable assumption to make about an older woman who comes into a mental health clinic for treatment for depression is that she cares primarily 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, men are more likely than women to be very sensitive to treatment costs and side effects.

Having a new patient answer the set of seven attitudinal questions dramatically increases the probability of identifying to which preference class the individual belongs. For illustrative purposes, consider the case of older women. These individuals have high unconditional probabilities of belonging to the Least 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 Least Sensitive class range from 0 to 100%. For example, one woman in this class answered that all attributes were *Very Important.* Based on these answers, this woman has a conditional probability of 100% of being in the More Sensitive class. Similarly, another woman who answered that effectiveness and the weight gain side effect were *Very Important* while all other attributes were *Somewhat/Pretty Important* has a 99% probability of being in the Somewhat Sensitive group. 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.

The three-class results were cross-validated by comparing model results across subsamples of the original data. We randomly selected 309 samples of 90 observations each from the original sample: 14 observations were randomly dropped from the original sample each time. We performed latent class analysis on each new subsample and determined whether individuals were assigned to the Least Sensitive, Most Sensitive, or Somewhat Sensitive groups. The degree of similarity between the grouping on the original data and the subsamples of data was calculated using the Hubert-Arabie Adjusted Rand Index (Hubert and Arabie 1985).[6] We then calculated an average Adjusted Rand Index by averaging over all the samples. The average adjusted Rand Index for the three-class latent-class model was 0.92, indicating that on average individuals were very likely to be assigned to the same group, even with small changes to the sample. This suggests that the latent class results are robust.

At the suggestion of a referee, we compared the latent class results with the results of a cluster analysis. Because we were interested in comparing the results of the two methods, we imposed three clusters rather than attempting to find the optimal number of clusters. Results were obtained using PROC CLUST (Ward's method) in SAS (SAS Institute Inc 1987), based on a proximity matrix calculated using Gower's dissimilarity coefficient. Table 3 shows the mean answers to each of the attitudinal questions for each cluster.

Cluster analysis supports the results from the latent-class model and indicates that the three groups differ in their sensitivity.[7] Ranking of the means in Table 3 within groups and comparision of means across groups shows that the characterization of the three groups is essentially the same. The primary difference between the two methods is that for cluster analysis, the ranked means within a group are significantly different from each other a smaller share of the time. The size of the three groups also differ somewhat between the two methods.

[6] The Adjusted Hubert-Arabie Rand Index is considered to be the best measure for measuring the level of agreement between two data partitions (Milligan and Cooper 1986). A number closer to 1 indicates a greater degree of agreement in the grouping allocations between the samples.

[7] The cluster analysis results were cross-validated using the average Adjusted Rand Index from 255 random samples of 90 observations each. The average adjusted Rand Index is 0.67, again indicating that on average individuals are likely to be assigned to the same group, even with small changes to the sample.

Discussion

This paper presented a cross-validated 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 varied among depressed patients. Three classes were identified that differed in their sensitivity to treatment costs and side effects. One class (Least Sensitive) cared primarily about treatment effectiveness; side effects and the cost of treatment had little impact on this class's treatment decisions. Another class (Most Sensitive) 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. A third class (Somewhat Sensitive) was in between the other two in its senstivity to treatment factors. Group membership is a function of age and gender; younger or male patients were more likely to be sensitive to treatment costs and side effects.

Latent class analysis is less well known than other grouping methods such as cluster analysis. Our results show that the general interpretation of the three groups using both techniques was fairly similar based on analysis of mean responses. Given the fact that clustering results can often vary as a function of the particular clustering technique employed, this is a good result but does not necessarily generalize to other datasets.

Although the interpretation of the cluster and latent class groupings seemed fairly similar, there are a number of arguments to be made for using latent class analysis rather than defaulting to cluster analysis. First, consider the approach of the two methods. 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, cluster group membership is typically assumed deterministic; individuals are allocated to groups as if the researcher knows with 100% certainty to which group individuals belong.

This is a strong assumption and one that in many cases may be wrong. This assumption can be loosened by using other techniques such as fuzzy clustering, but these methods are fairly complex.

Software is not a reason to default to cluster analysis. There are many statistical packages widely available to do cluster analysis; however, knowing which of the many techniques and algorithms is most appropriate to employ involves a steep learning curve. There are now a number of both free and proprietary software that can be used to do latent class analysis. Free software such as LEM is very efficient and fast but requires some background understanding to apply. Proprietary software such as Latent GOLD (Vermunt and Magidson 2000) and Mplus (Muthen and Muthen 1998-2004) are increasingly making latent class analysis easily accessible.[8]

A third argument to be made on behalf of latent class analysis is the type of output produced. Cluster analysis essentially produces mean responses that allow you to characterize the groups and identify the individuals that belong to each group. Latent class analysis can produce this; it also produces response probabilities and conditional probabities. The response probabilities enhance the ability to characterize the groups and can provide guidance for clinicians about how treatment preferences vary among demographic groups. Because they are probabilities, researchers can generate estimates of the size of these classes for different patient populations. The conditional probabilities also appear to have a number of practical applications. For example, 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 with a high probability. 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

[8] See http://ourworld.compuserve.com/homepages/jsubersax/soft.htm for a brief description and links to a number of latent class software sites.

information about patient preferences should allow the clinician to be of greater assistance in helping the patient to pick the best treatment plan.

More study is required on the differences between latent class and cluster analysis. However, we think this paper demonstrates a number of reasons why researchers should consider using latent class analysis. By applying the latent-class model, investigators can identify distinct classes that differ in their treatment preferences and concerns. 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 clincians who recruited them, in particular Bonnie Mucklow, Stew Nyholm, John Romeo, Nancy Rogers, and Marsha Visscher. Thanks to Bill Breffle, Alok Bohara, Roger von Haefen, Nick Flores, Don Waldman, and three anonymous reviewers for valuable comments and suggestions.

References

Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716–723.

Aptech Systems I. 1995. GAUSS, volume 3.2. Maple Valley, WA.

Bandeen-Roche K, Miglioretti D, Zeger S, Rathouz P. 1997. Latent variable regression for multiple discrete outcomes. J Am Stat Assoc 92:1375–1386.

Bartholomew D, Knott M. 1999. Latent Variable Models and Factor Analysis. New York: Oxford University Press Inc., second edition, 214 pp.

Bozdogan H. 1987. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika 52:345–370.

Brundage M, Feldman-Stewart D, Cosby R, Gregg R, Dixon P, Youssef Y, Mackillop W. 2001. Cancer patients' attitudes toward treatment options for advanced non-small cell lung cancer: implications for patient education and decision support. Patient Educ Couns 45:149–157.

Clogg C, Goodman L. 1984. Latent structure analysis of a set of multidimensional contingency tables. J Am Stat Assoc 79:762–771.

Cooper L, Brown C, Vu H, Palenchar D, Gonzales J, Ford D, Powe N. 2000. Primary care patients' opinions regarding the importance of various aspects of care for depression. Gen Hosp Psych 22:163–173.

De Menezes L, Bartholomew D. 1996. New developments in latent structure analysis applied to social attitudes. Jo R Stat Soc: Ser A 159:213–224.

Dempster A, Laird N, Rubin D. 1977. Maximum likelihood from incomplete observations. Jo R Stat Soc: Ser B 39:1–38.

Demyttenaere K. 1997. Compliance during treatment with antidepressants. J Affect Disord 43:27–39.

Dwight-Johnson M, Sherbourne C, Liao D, Wells K. 2000. Treatment preferences among depressed primary care patients. J Gen Intern Med 15:527–534.

Eid M, Langeheine R, Diener E. 2003. Comparing typological structures across cultures by multigroup latent class analysis - a primer. J Cross Cult Psych 34:195–210.

Forman A. 2003. Latent class model diagnostics - a review and some proposals. Computational Statistics and Data Analysis 41:549–559.

Green C, Brazier J, Deverill M. 2000. Valuing health-related quality of life: a review of health state valuation techniques. Pharmacoeconomics 2000:151–165.

Hubert L, Arabie P. 1985. Comparing partitions. Journal of Classification 2:193–218.

Hurvich M, Tsai C. 1989. Regression and time series model selection in small samples. Biometrika 76:297–307.

Kremer T, Gesten E. 2003. Managed mental health care: the client's perspective. Prof Psych 34:187–196.

Lin E, VonKorrff M, Katon W, Bush T, Simon G, Walker E, Robinson P. 1995. The role of the primary care physician in patients' adherence to antidepressant therapy. Med Care 33:67–74.

McCombs J, Nicho M, Stimmel G, Sclar D, Beasley C, Gross L. 1990. The cost of antidepressant drug therapy failure: a study of antidepressant use patterns in a Medicaid population. J Clin Psychiatry 51:60–69.

McCutcheon A. 1987. Sexual morality, pro-life values, and attitudes toward abortion - a simultaneous latent structure analysis for 1978-1983. Socio Methods Res 16:256–275.

McCutcheon A, Nawojcyzk M. 1995. Making the break - popular sentiment toward legalized abortion among American and Polish Catholic laities. International Journal of Public Opinion Research 7:232–252.

Milligan G, Cooper M. 1986. A study of the comparability of external criteria for hierchical cluster analysis. Multivariate Behavioral Research pp. 441–458.

Montgomery A, Harding J, Fahey T. 2001. Shared decision making in hypertension: the impact of patient preferences on treatment choice. Fam Pract 18:309–313.

Morey E, Thacher J, Breffle W. 2005. Using angler characteristics and attitudinal data to identify environmental preference classes: A latent-class model. Environmental and Resource Economics (in press) .

Muthen B, Muthen L. 1998-2004. Mplus User's Guide. Los Angeles, CA: Muthen and Muthen.

Myers E, Branthwaite A. 1992. Out-patient compliance with antidepressant medication. Br J Psychiatry 160:83–86.

Nease R, Kneeland T, O'Connor G, Sumner W, Lumpkins C, Shaw L, Pryor D, Sox H. 1995. Variation in patient utilities for outcomes of the management of chronic stable angina: implications for clinical practice guidelines. J Am Med Assoc 273:1185–1190.

O'Brien B, Novosel S, Torrance G, Streiner D. 1995. Assessing the economic value of a new anti-depressant: a willingness-to-pay approach. Pharmacoeconomics 8:34–45.

Revicki D, Wood M. 1998. Patient-assigned health state utilities for depression-related outcomes: differences by depression severity and antidepressant medications. J Affect Disord 48:25–36.

SAS Institute Inc. 1987. SAS User's Guide 8.02. Cary, NC: SAS Institute Inc.

Schwarz G. 1978. Estimating the dimension of a model. The Annals of Statistics 6:461–464.

Sestoft D, Pedersen L, Bendsen B, Hancke B, Larsen A, Mikkelsen S, Unden M, Kreiner S. 1998. The effect of electroconvulsive therapy on patient's attitudes to treatment of depression. Nord J Psychiatry 52:31–37.

Simon G, Von Korrff M, Heiligenstein J, Revicki D, Grothaus L, Katon W, Wagner E. 1996. Initial antidepressant choice in primary care: effectiveness and cost of fluoxetine vs. tricyclic antidepressants. J Am Med Assoc 275:1897–1902.

Thompson D, Bueseching D, Grogor K, Oster G. 1996. Patterns of anti-depressant use and their relation to costs of care. Am J Man Care 2:1239–1246.

Titterington D, Smith A, Makov U. 1985. Statistical Analysis of Finite Mixture Distribution. New York: John Wiley & Sons, first edition.

Tsevat J, Dawson N, Wu A, Lynn J, Soukup J, Cook E, Vidaillet H, Philips R. 1998. Health values of hospitalized patients 80 years or older. J Am Med Assoc 279:371–375.

Vermunt J. 1997. LEM 1.0: A general program for the analysis of categorical data. Tilburg: Tilburg University.

Vermunt J, Magidson J. 2000. Latent GOLD User's Manual. Belmont, MA: Statistical Innovations Inc, 185 pp.

von Daiver M. 1997. Bootstrapping goodness-of-fit statistics for sparse categorical data: Results of a Monte Carlo study. Methods of Pyschological Research Online 2.

Walburn J, Gray R, Gournay K, Quraishi S, David A. 2001. Systematic review of patient and nurse attitudes to depot antipsychotic medication. Br J Psychiatry 179:300–307.

Wedel M, Kamakura W. 2000. Market segmentation: conceptual and methodological foundations. Boston: Kluwer Academic Publishers, second edition.

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.

Wu Z, Freeman J, Greer A, Freeman D, Goodwin J. 2001. The influence of patients' concerns on surgeons' recommendations for early breast cancer. Eur J Cancer Care 10:100–106.

Yamaguchi K. 2000. Multinomial logit latent-class regression models: an analysis of the predictors of gender-role attitudes among Japanese women. American Journal of Sociology 105:1702–1740.

Table 1: Goodness of Fit Measures for Latent Class Model

Model	**lnL**	**AIC**	**CAIC**	**AIC$_C$**	**BIC**	**Pearson P-value**	**Read-Cressie P-value**
1 group	−698.24	1424.48	1461.64	1429.93	730.75	0.08	0.01
2 groups	−649.20	1350.41	1419.41	1370.30	709.58	0.16	0.04
3 groups	−617.33	1306.65	1402.19	1349.26	700.92	0.27	0.13

Table 2: Three Classes: Predicted Response Probabilities and Standard Errors

		"Most"	"Least"	"Some"
Anti-Depressants	Not at All/Not Very	6% (.002)	20% (.006)	22% (.003)
	Somewhat/Pretty	54% (.009)	54% (.011)	59% (.004)
	Very	41% (.008)	27% (.007)	19% (.003)
Effectiveness	Not at All/Not Very	0% (.)	0% (.)	2% (.000)
	Somewhat/Pretty	27% (.007)	7% (.003)	30% (.005)
	Very	73% (.007)	93% (.003)	68% (.005)
Therapy Hours	Not at All/Not Very	24% (.007)	27% (.008)	17% (.003)
	Somewhat/Pretty	48% (.009)	57% (.010)	71% (.004)
	Very	28% (.008)	17% (.005)	12% (.002)
Cost	Not at All/Not Very	12% (.004)	43% (.008)	15% (.003)
	Somewhat/Pretty	48% (.008)	40% (.010)	62% (.005)
	Very	40% (.008)	17% (.006)	24% (.004)
Weight-Gain Side Effect	Not at All/Not Very	9% (.003)	23% (.006)	12% (.002)
	Somewhat/Pretty	14% (.004)	53% (.009)	49% (.005)
	Very	77% (.007)	24% (.007)	39% (.005)
Sex-Drive Side Effect	Not at All/Not Very	0% (.)	82% (.006)	0% (.)
	Somewhat/Pretty	0% (.)	14% (.005)	83% (.003)
	Very	100% (.)	4% (.002)	17% (.003)
No-Orgasm Side Effect	Not at All/Not Very	3% (.001)	92% (.003)	0% (.)
	Somewhat/Pretty	0% (.)	8% (.003)	83% (.003)
	Very	97% (.001)	0% (.)	17%(.003)

Table 3: Means by Group: Cluster and Latent Class Analysis

	Latent Class Analysis			**Cluster Analysis**		
Importance	**"Most"**	**"Least"**	**"Some"**	**"Most"**	**"Least"**	**"Some"**
Anti-Depressants	2.33	2.07	1.98	2.30	1.96	2.03
Effectiveness	2.73	2.93	2.66	2.70	2.93	2.69
Therapy Hours	2.03	1.90	1.95	2.08	1.71	2.03
Cost	2.27	1.73	2.10	2.33	1.50	2.17
Weight-Gain	2.67	2.00	2.27	2.50	1.89	2.44
Sex-Drive	3.00	1.20	2.15	2.88	1.25	2.03
No-Orgasm	2.94	1.07	2.15	2.83	1.18	1.97
Group Size	33	30	41	40	28	36

1=Not at All/Not Very Important, 2=Somewhat/Pretty Important, 3=Very Important.

Table 4: Young and Male are More Likely to be Sensitive to Treatment Costs and Side-Effects

	Unconditional Membership Probabilities, Standard Errors, and Group Sizes			
	Female		**Male**	
	18-40	**40+**	**18-40**	**40+**
"Most Sensitive" Group	35% (.009)	21% (.008)	55% (.011)	0% (.)
	($n = 15$)	($n = 8$)	($n = 10$)	($n = 0$)
"Least Sensitive" Group	20% (.006)	52% (.011)	0% (.)	34% (.008)
	($n = 8$)	($n = 19$)	($n = 0$)	($n = 3$)
"Somewhat Sensitive" Group	45% (.010)	27% (.008)	45% (.011)	66% (.008)
	($n = 18$)	($n = 10$)	($n = 7$)	($n = 6$)