J Fam Econ Iss (2015) 36:451–461
DOI 10.1007/s10834-014-9407-2

ORIGINAL PAPER

The Method of Generalized Ordered Probit with Selectivity: Application to Marital Happiness

Karen Leppel

Published online: 24 May 2014

Abstract This paper introduces and demonstrates the use of the method of Generalized Ordered Probit with Selectivity (GOPS). It simultaneously estimates (1) functions that describe the relations between the explanatory variables and an ordered dependent variable such as marital happiness, and (2) the selectivity function that pertains to being married. It does not presume the equality-of-slopes assumption of conventional ordered probit. The GOPS method is demonstrated for modeling marital happiness. The results are compared to those from conventional ordered probit and generalized ordered probit without selectivity.

Keywords Ordered probit · Selectivity · Equality-of-slopes assumption · Marital happiness

Introduction

Over the past few decades, researchers have frequently worked with categorical dependent variables. The simplest of the methods, binary probit, is used for dependent variables with only two possible values. For multi-level dependent variables, ordered probit was designed. To extend ordered probit to circumstances in which the impact of the explanatory variables differs with the level of the dependent variable, generalized ordered probit was developed (Greene and Hensher 2008, pp. 79–80).

In some situations, selectivity is an issue. For example, when a researcher wants to estimate an equation in which the dependent variable is marital satisfaction, such data can only be provided for individuals who are married. When the dependent variable is only available for a specific subset of the population, such as married individuals, and unobservable characteristics of that subset are related to the value of the dependent variable, the coefficients of the explanatory variables are biased. To handle situations with a continuous dependent variable that is subject to selectivity issues, Heckman (1976) developed a method that combined the use of binary probit with least squares regression. While numerous researchers have explored marital quality using least squares regression and approximately continuous variables based on composite measures[1] (Maume and Sebastian 2012; Pedersen and Minnotte 2012; Tsang et al. 2003; Van den Troost et al. 2006), they have not used Heckman's method to correct for selectivity.

When the dependent variable is a discrete, ordered one, Heckman's method cannot be used. For that type of situation, we need a new method that can (1) handle a multi-level dependent variable and (2) deal with selectivity issues. That method is Generalized Ordered Probit with Selectivity (GOPS). This study explains and develops the method and demonstrates its use for modeling marital happiness. In essence, working with a sample of both married and unmarried individuals, the procedure simultaneously estimates a function for determining whether an individual is likely to be married along with a separate function for examining the level of marital happiness. The paper begins by describing the selectivity problem in more detail.

K. Leppel (✉)
School of Business Administration, Widener University,
One University Place, Chester, PA 19013, USA
e-mail: kleppel@widener.edu

[1] For their dependent variable, Maume and Sebastian (2012) used the mean of 15 five-point Likert-scale items that assessed various aspects of a respondent's relationship. Pedersen and Minnotte (2012) used the sum of seven relationship items to create a score ranging from 7 to 35. Tsang et al. (2003) used 11 items resulting in a scale range from 11 to 34. Van den Troost et al. (2006) used a measure based on the sum of seven seven-point Likert-scale items.

The Selectivity Problem

In the case of marital satisfaction, sample selectivity problems happen because we only have data on marital satisfaction for individuals who are married. Suppose there are certain observed characteristics W that increase an individual's likelihood of marriage. There are, however, some individuals who marry (and therefore get into our married sample) despite being low on those observed characteristics; they marry because they have high levels of some unobserved characteristic V that offset what they lack in W.

For example, suppose the unobservable V is the importance an individual places on marriage. Some people are low in the observed W but have unusually high values of V and work unusually hard at making their marriages successful. They then may experience high levels of marital satisfaction despite being low on the observed W. In this case, we would underestimate the effects of W.

As another example, suppose the unobservable V is an inability or unwillingness to live independently or a desire to conform to the social convention of marriage. Some people are low in the observed W but have unusually high values of dependency/conformity V. If a dependent or conforming nature V tends to make a weak or unhappy marriage partner, high-V people may have lower levels of marital satisfaction. They then may experience low marital satisfaction not only because they are low on the observed characteristics W but also because they are high on dependency/conformity V. In this case, we would overestimate the effects of the observed W. Another possible unobservable factor of this type could be charisma or "sex appeal." Individuals low on observed W may marry because they have unusually high levels of a characteristic that makes them extremely attractive to potential marriage partners. After marriage, however, this characteristic could cause a great deal of jealousy and conflict, lowering marital satisfaction. So as before, these individuals may experience low marital satisfaction not only because they are low on the observed characteristics W but also because they are high on charisma V, and once again, we would overestimate the effects of the observed W.

We see here that whether the unobservable factor V is positively or negatively associated with marital satisfaction, coefficient estimates for the observed factors W may be biased by a failure to take sample selectivity into account. While we can discuss these unobservable factors in theory, in practice, the situation can be quite complicated. There may be multiple unobservable factors which work in different directions; some may have a positive effect on marital satisfaction, some a negative effect. In addition, some of these factors may be strongly correlated with certain explanatory variables and therefore have a bigger biasing effect on those variables.

Since we generally do not know the unobservable factors, it is difficult to anticipate the specific effects of the sample selectivity bias. It is possible, nonetheless, to adjust for possible bias. Before explaining the method of doing so, the paper discusses the conventional ordered probit model, explains the deficiencies of that model, and modifies it to develop the generalized ordered probit technique without selectivity. The analysis then corrects for the shortcomings of that method by adapting it to develop the generalized ordered probit procedure with selectivity.

Mathematical Development

Conventional Ordered Probit Model

Consider, for example, the usual ordered probit model in a situation in which the observed dependent variable (*S*) has three categories. The model is based on a latent satisfaction variable (S^*) which is a function of the independent variables (*X*): $S^* = X\beta_1 + \varepsilon$, where β_1 is a vector of unknown parameters, c is a constant threshold, and ε is a stochastic disturbance term. It is assumed that

$$S = \begin{cases} 1 & \text{if} \quad S^* \leq 0 \quad \text{or} \quad \varepsilon \leq -X\beta_1 \\ 2 & \text{if} \quad 0 < S^* \leq c \quad \text{or} \quad -X\beta_1 < \varepsilon \leq -X\beta_1 + c \\ 3 & \text{if} \quad S^* > c \quad \text{or} \quad \varepsilon > -X\beta_1 + c \end{cases}$$

The existence of a single latent variable function and a constant threshold implies an equality-of-slopes assumption. In other words, the impact of the explanatory variables is the same regardless of the satisfaction level.

Generalized Ordered Probit Model

When the equality-of-slopes assumption does not hold, a generalized ordered probit technique is needed. The assumption is avoided by allowing the latent variable's threshold between categories to be a function of the independent variables (Greene and Hensher 2008, pp. 79–80). It is still assumed that $S^* = X\beta_1 + \varepsilon$. However, it is also assumed that τ is a vector of unknown parameters, and

$$S = \begin{cases} 1 & \text{if} \quad S^* \leq 0 \\ 2 & \text{if} \quad 0 < S^* \leq X\tau \\ 3 & \text{if} \quad S^* > X\tau \end{cases}$$

Equivalently, letting $\beta_2 = \beta_1 - \tau$,

$$S = \begin{cases} 1 & \text{if} \quad \varepsilon \leq -X\beta_1 \\ 2 & \text{if} \quad -X\beta_1 < \varepsilon \leq -X\beta_2 \\ 3 & \text{if} \quad \varepsilon > -X\beta_2 \end{cases}$$

In this model, the impact of the explanatory variables varies with the satisfaction level. Whether an individual is in satisfaction level 2 rather than satisfaction level 1 depends on the function $X\beta_1$, while whether an individual is in satisfaction level 3 rather than satisfaction level 2 depends on the function $X\beta_2$. However, additional problems may result from sample selection issues.

Generalized Ordered Probit Model with Selectivity

We can allow and correct for the possibility of sample selection problems using the technique described below. In addition to estimating the parameters pertaining to marital satisfaction, we need to estimate a function for determining whether an individual is likely to be married. To do so, an additional latent variable is incorporated into the generalized ordered probit model. Individuals are married if the utility of marriage exceeds the utility of remaining unmarried. The difference in the utilities (Y^*) is unobservable, but assumed to be a function of personal characteristics (Z). That is, $Y^* = Z\gamma + u$, where γ is a vector of unknown parameters and u is a stochastic disturbance term. An individual is married ($M = 1$) if $Y^* > 0$ or $u > -Z\gamma$. An individual is not married ($M = 0$) if $Y^* \leq 0$ or $u \leq -Z\gamma$. The ε in the latent marital satisfaction function and u are assumed to have a bivariate normal distribution with correlation ρ.

The impacts of the unobservable V-factors get absorbed into the error terms ε and u. A positive correlation ρ indicates that the V-factors that increase the probability of marriage are positively associated with marital satisfaction. A negative ρ indicates that V-factors that increase the probability of marriage are negatively associated with marital satisfaction. If there are multiple factors operating in different directions, the sign of ρ tells us the predominant direction. If ρ is not significantly different from zero, then sample selection bias is not a problem and the estimation results should be approximately the same for generalized ordered probit with or without selectivity.

To estimate the GOPS model, we need to calculate the probabilities of the different marital satisfaction levels. Let Φ_1 denote a univariate cumulative normal distribution and Φ_2 a bivariate cumulative normal distribution. The probabilities of the various cases are given below. The details of the calculations (for the general situation with s satisfaction levels) are provided in the appendix.

$$\begin{aligned}&\text{Case 1}: \Pr(M = 1 \text{ and } S = 1)\\&= \Pr(u > -Z\gamma \text{ and } \varepsilon \leq -X\beta_1)\end{aligned}$$

$$\begin{aligned}&\text{Case 2}: \Pr(M = 1 \text{ and } S = 2)\\&= \Pr(u > -Z\gamma \text{ and } -X\beta_1 \leq \varepsilon \leq -X\beta_2)\end{aligned}$$

$$\begin{aligned}&\text{Case 3}: \Pr(M = 1 \text{ and } S = 3)\\&= \Pr(u > -Z\gamma \text{ and } \varepsilon \geq -X\beta_2)\end{aligned}$$

$$\text{Case 4}: \Pr(M = 0) = \Pr(u \leq -Z\gamma) = \Phi_1(-Z\gamma)$$

The probabilities of the four possibilities can then be used to construct a (weighted) log likelihood function, which can be estimated using a maximum likelihood technique such as the SAS NLMIXED procedure. (Sample SAS programming statements are available from the author.) After discussing the data used, the results of the application of the method of Generalized Ordered Probit with Selectivity to marital happiness will be examined.

Data

The data used in this study (Smith et al. 2011) came from the 1972-2010 cumulative file of the General Social Survey (GSS). The National Opinion Research Center (NORC) designed the GSS to collect data on demographic characteristics and attitudes of US residents. The survey is limited to individuals 18 years of age or older and asks respondents a set of core questions every year, while it includes topical modules of questions on a rotating basis. The GSS provides survey weights which were used in the current analysis. The survey included a question on marital happiness: "Taking all things together, how would you describe your marriage? Would you say that your marriage is very happy, pretty happy, or not too happy?" (Smith et al., p. 327).

In addition to gender and marital status, the GSS contains information about personal characteristics that might be related to marital happiness: age, race, social class (upper or lower versus middle or working class),[2] number of children in the household, and hours worked by respondent and respondent's spouse. The data also contain information that enables the researcher to compare characteristics of the respondent and spouse. Using that data, variables were created for same educational level of attainment and same religion group. Levels of educational attainment were: less than high school diploma, high school diploma, junior college degree, bachelor's degree, and graduate degree. Religion groups were: Catholic, non-

[2] The measure of respondent's real income provided in the cumulative GSS data file has severe limitations. First, 41 % of the observations in that file have no data on income (defined as annual earnings from one's reported occupation). Also, NORC has imputed this measure from categorical data for seven separate income variables used at different points in history, based on the midpoints of income brackets whose definitions varied over time; using such imputations in our analysis could introduce problematic measurement errors. Moreover, an individual's perception of relative earnings may be as important as or more important than actual earnings in influencing happiness. Therefore, the respondent's assessment of social class was used instead of the GSS income variable in the current analysis of marital happiness. The survey question was phrased, "If you were asked to use one of four names for your social class, which would you say you belong in: the lower class, the working class, the middle class, or the upper class?" (Smith et al. 2011, p. 383).

Catholic Christian, Jewish, other religion, and no religion. Traditional roles of husband as breadwinner and wife as homemaker suggest that the difference between the occupational statuses of spouses may influence marital happiness. Therefore, the SEI gap or difference between the Socioeconomic Index (SEI) of the respondent's occupation and the SEI of the respondent's spouse's occupation, as provided by the GSS, was also included in the analysis.[3,4]

For the equation specifying whether the respondent was married, variables included age, white race, whether respondent resided in the southern US, whether respondent had a bachelor's degree, and four dummy variables for religion: Catholic, non-Catholic Christian, Jewish, and other religion. The reference category for religion was no religion.

Random samples of 4500 men and 4500 women were taken who were either married and had complete data on the variables in both the marital status and marital happiness equations, or who were not married and had complete data for the marital status equation. The phrase "full samples" is used to denote the 4500 men and 4500 women. These samples contained 1478 married men, 1169 married women, 3022 unmarried men, and 3331 unmarried women. The definitions of the variables are summarized in Table 1.

Table 1 Definitions of variables

Variable	Definition
Marital happiness	3 If respondent specifies that his/her marriage is very happy; 2 if pretty happy; 1 if not too happy
Age	Age in years
White	1 if race of respondent is white; 0 otherwise
Number of children in household	Number of children living in the household
Lower class[a]	1 if respondent self-identifies as lower class; 0 otherwise
Upper class[a]	1 if respondent self-identifies as upper class; 0 otherwise
Hours worked	Hours worked last week by respondent
Spouse's hours worked	Hours worked last week by respondent's spouse
Same educational level	1 if spouse has same educational level as respondent (levels are: less than high school diploma, high school diploma, junior college degree, bachelor's degree, graduate degree
Same religion group	1 if spouse is same religion group as respondent (groups are: Catholic, non-Catholic Christian, Jewish, other religion, or no religion); 0 otherwise
Socioeconomic index gap	Socioeconomic index (SEI) of respondent's occupation minus socioeconomic index of spouse's occupation (SEI values are between 0 and 100; so SEI Gap is between -100 and 100.) The SEI value is set equal to zero for individuals who are not working.
Additional variables used only in Marital Status equation	
Southern residence	1 if respondent lives in the Southern region of the US; 0 otherwise
Bachelor's degree	1 if respondent has a bachelor's degree; 0 otherwise
Catholic[b]	1 if respondent is Catholic; 0 otherwise
Non-Catholic Christian[b]	1 if respondent is non-Catholic Christian; 0 otherwise
Jewish[b]	1 if respondent is Jewish; 0 otherwise
Other religion[b]	1 if respondent has a religion other than Catholic, non-Catholic Christian, or Jewish

[a] The social class reference category is middle or working class

[b] The religion reference category is no religion

Estimation Results

Results of the GOPS Procedure

Preliminary analysis showed significant differences in the level of marital happiness for men versus women (Table 2), with the men happier than the women. Furthermore, a likelihood ratio test indicated a much better fit when the systems were estimated separately for men and women, rather than together ($\chi^2 = 257.6$, df $= 35$, $p < 0.0001$). Therefore, this paper shows the results of the separate systems. The estimated equations predict marital happiness outcomes correctly for 65.1 % of the married men and 61.8 % of the married women.[5,6]

Table 3 provides descriptive statistics for married respondents for the explanatory variables in the marital happiness equations, while Table 4 shows the estimated

[3] The Socioeconomic Index was created by Keiko Nakao and Judith Treas, based on the 1989 GSS study of occupational prestige and procedures developed by Otis Dudley Duncan. It captures occupational status through a "description of the socioeconomic hierarchy of occupations in the intergenerational transmission of status" (Nakao and Treas 1992, p. 3).

[4] While the GSS contains information about the respondent's self-assessed health, that variable was not used here, due to its high incidence of missing values. The GSS does not provide information on the race of the respondent's spouse. Attempts to explore linear trends by including the survey year resulted in multicollinearity problems.

[5] These percentages are comparable to those for the conventional ordered probit (65.2 for men and 60.5 for women) and the generalized ordered probit without selectivity (65.4 for men and 61.4 for women).

[6] The correlations between the error terms were found to be statistically significantly negative for both men and women. This finding suggests that random factors that increase the probability of marriage tend to decrease marital happiness. These factors result in higher probabilities of marriage but lower levels of marital happiness.

coefficients and associated p values for each variable in those equations. (Table 5 provides descriptive statistics for the full sample for the variables in the marital status equation, while Table 6 shows the estimated coefficients and associated p values for each variable in that equation.) A positive coefficient in the function $X\beta_1$ indicates that an increase in the specified variable increases the probability of being pretty happy versus not too happy. A positive coefficient in the function $X\beta_2$ indicates that an increase in the specified variable increases the probability of being very happy versus pretty happy.

To facilitate interpretation of the results, the marginal effects of the explanatory variables on the marital happiness probabilities are also shown in Table 4. The calculated probabilities were conditioned on being married. Base case characteristics are for a respondent who is age 50 and of white race, has no children living in the household, self-identifies as middle or working class, works 40 h per week, has a spouse who also works 40 h per week, has the same educational level and the same religion group as the spouse, and a socioeconomic index gap of zero. The respondent is a non-Catholic Christian, does not have bachelor's degree, and does not reside in the south. For dummy variables, the marginal effect is calculated as (the conditional probability of specified marital happiness level with $X = 1$ and other base case values as described) minus (the conditional probability of specified marital happiness level with $X = 0$ and other base case values as described). The marginal effect of age is calculated as (the conditional probability of the specified marital happiness level evaluated at age 51 and other base case values as given) minus (the conditional probability evaluated at age 50 and other base case values as given). The marginal effect of hours worked (for respondent or for spouse) is calculated as (the conditional probability of the specified marital happiness level evaluated with hours worked = 41 and other base case values as given) minus (the conditional probability evaluated with hours worked = 40 and other base case values as given).

The results of the GOPS estimation of marital happiness are not directly comparable to those of previous researchers. Not only have other researchers used different measures for their dependent variable, but they have also included different explanatory variables. In the discussion that follows, however, it is indicated whether the findings are consistent with those of other studies.

For both men and women, the signs of the coefficients of the age and age-squared variables suggested that as age increases, the probability of being very happy versus pretty happy with one's marriage and the probability of being pretty happy versus not too happy initially decreased and later increased. The marginal effects of age indicate that a 51-year-old man with other characteristics as in the base

Table 2 Percent in marital happiness categories by sex

Sex	Not too happy	Pretty happy	Very happy
Males	2.0	32.9	65.1
Females	3.4	36.3	60.3

A Chi squared test concluded that marital happiness and sex of respondent are not independent variables ($\chi^2 = 9.79$, df = 2, $p = 0.0075$)

Table 3 Summary statistics for variables in marital happiness equations for married respondents

Characteristic	Males ($n = 1478$)	Females ($n = 1169$)
Percentages		
White	0.8701	0.8674
Lower class[a]	0.0433	0.0522
Upper class[a]	0.0298	0.0180
Same educational level	0.5670	0.5885
Same religion group	0.8092	0.8050
Means (standard deviations)		
Age	51.4899	49.4294
	(16.5029)	(16.2061)
Number of children in household	0.7652	0.8015
	(1.1471)	(1.1929)
Hours worked per week by respondent	27.4181	16.9564
	(25.4049)	(20.8533)
Hours worked by spouse	16.1455	25.9204
	(20.3419)	(24.5704)
Socioeconomic index gap	7.9882	−5.1419
	(29.7652)	(29.8046)

[a] The social class reference category is middle or working class

case would be expected to have a probability of being very happy that is 0.0021 less than an otherwise comparable 50-year-old man, while a 51-year-old woman would be expected to have a probability of being very happy that is 0.0029 less than an otherwise comparable 50-year-old woman. Maume and Sebastian (2012) included age but not age-squared in their regressions and found no statistically significant effects of age on marital quality.

Also according to Table 4, white men appeared to be happier with their marriages than non-white men were. In particular, the coefficient in the $X\beta_1$ function suggests that the probability of being pretty happy versus not too happy is greater for white men than for comparable non-white men. The probability of being very happy with one's marriage versus pretty happy is greater for white men than for non-white men, but not significantly so. The probability of being not too happy is 0.0228 lower for white men than for similar non-white men. Maume and Sebastian (2012)

Table 4 GOPS estimation results for marital happiness

Variable	Function $X\beta_1$ Coefficient	p value	Function $X\beta_2$ Coefficient	p value	Marginal Effects on Marital Happiness Probabilities[a] Not too happy	Pretty happy	Very happy
Males ($n = 1478$)							
Constant	4.5227***	<0.0001	2.6531**	0.0025			
Age	−0.0936**	0.0079	−0.0733**	0.0012	0.0004	0.0017	−0.0021
Age squared	0.0008*	0.0181	0.0006**	0.0018			
White	0.3154†	0.0663	0.1915	0.1414	−0.0228	−0.1142	0.1370
Number of children in household	−0.0248	0.7123	−0.0490†	0.0650	0.0009	0.0189	−0.0198
Lower class	−0.3041	0.2325	−0.0273	0.8358	0.0152	−0.0042	−0.0109
Upper class	−0.1240	0.7576	0.4180*	0.0343	0.0050	−0.1479	0.1429
Hours worked	0.0124*	0.0214	−0.0012	0.4882	−0.0004	0.0009	−0.0005
Spouse's hours worked	−0.0091	0.1146	0.0007	0.7473	0.0003	−0.0006	0.0003
Same educational level	0.0494	0.7274	0.0144	0.7967	−0.0018	−0.0039	0.0058
Same religion group	0.1105	0.5156	0.1171	0.1385	−0.0044	−0.0437	0.0481
Socioeconomic index gap	−0.0041	0.2905	0.0021	0.1314	0.0001	−0.0010	0.0008
Females ($n = 1169$)							
Constant	4.4850***	<0.0001	3.1403***	<0.0001			
Age	−0.0966***	0.0008	−0.0897***	<0.0001	0.0000	0.0029	−0.0029
Age squared	0.0009 **	0.0011	0.0008***	<0.0001			
White	−0.0401	0.8107	−0.0244	0.8724	−0.0077	−0.0791	0.0868
Number of children in household	−0.1599 **	0.0012	−0.1324***	0.0002	0.0102	0.0486	−0.0588
Lower class	−0.4006†	0.0673	−0.0475	0.7249	0.0332	−0.0124	−0.0208
Upper class	−0.1478	0.7166	0.1474	0.4945	0.0093	−0.0714	0.0621
Hours worked	−0.0031	0.4632	0.0005	0.8245	0.0002	−0.0004	0.0002
Spouse's hours worked	0.0046	0.1986	−0.0014	0.4374	−0.0002	0.0009	−0.0006
Same educational level	0.1802	0.1376	−0.1087†	0.0768	−0.0117	0.0579	−0.0462
Same religion group	0.0675	0.6425	0.2060*	0.0108	−0.0039	−0.0883	0.0921
Socioeconomic Index Gap	−0.0033	0.2408	−0.0020	0.1535	0.0002	0.0007	−0.0009

For males, the correlation between the marital happiness error term and the marital status error term is −0.6455 with $p = 0.0103$. For females, the correlation between the marital happiness error term and the marital status error term is −0.6186 with $p = 0.0136$. A positive coefficient in function $X\beta_1$ indicates that an increase in the specified variable increases the probability of being pretty happy versus not too happy. A positive coefficient in function $X\beta_2$ indicates that an increase in the specified variable increases the probability of being very happy versus pretty happy

[a] Probabilities shown were conditioned on being married. Base case characteristics are for a respondent who is age 50, of white race, has no children living in household, self-identifies as middle or working class, 40 h worked per week by respondent and respondent's spouse, same educational level and same religion group as spouse, and socioeconomic index gap of zero. The respondent is a non-Catholic Christian, does not have bachelor's degree, and does not reside in the south. For dummy variables, the marginal effect is calculated as (probability of specified marital happiness level with $X = 1$ and other base case values as described) minus (probability of specified marital happiness level with X = 0 and other base case values as described). The marginal effect of age is calculated as (probability of the specified marital happiness level evaluated at age 51 and other base case values as given) minus (probability evaluated at age 50 and other base case values as given). The marginal effect of hours worked (for respondent or for spouse) is calculated as (probability of the specified marital happiness level evaluated with hours worked = 41 and other base case values as given) minus (probability evaluated with hours worked = 40 and other base case values as given)

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

also reported evidence that white men have higher levels of marital quality than non-white men.

The results of the GOPS estimation of marital happiness for the other variables are discussed briefly below. The findings show that the number of children in the household was negatively associated with marital happiness, which is consistent with the results of Tsang et al. (2003). Pedersen and Minnotte (2012) found a significantly negative relation between the number of children and marital satisfaction in dual-earner couples in their correlation analysis, but not in their regressions in which they included variables to explore the impact of workplace culture on marital satisfaction.

Also, compared to middle- or working-class women, women who self-identify as lower class were less likely to be pretty happy versus not too happy with their marriages, while upper-class men were more likely to be very happy versus pretty happy. Insofar as the terms upper and lower class are subjective and relative, they may reflect satisfaction or dissatisfaction with perceived socioeconomic position in general and financial situation in particular; the results concerning class are thus consistent with the findings by Tsang et al. (2003) of a negative relation between dissatisfaction with the financial situation and marital happiness. Van den Troost et al. (2006) also reported evidence of a negative relation between income dissatisfaction and marital satisfaction among women.

Among men, working more hours was positively related to the probability of being pretty happy versus not too happy; the hours worked by their spouses appeared to have little association with marital happiness. Among women, their hours and those of their spouses appeared to have little relation to marital happiness. The positive effect of men's hours worked on their marital happiness may reflect the effect of financial stability. The lack of an effect of women's hours on their marital happiness may suggest that the number of hours is less important than the particular schedule's effect on work-family conflict, which Maume and Sebastian (2012) report influenced women's marital quality.

Table 5 Summary statistics for variables in marital status equation for full sample

Characteristic	Males ($n = 4500$)	Females ($n = 4500$)
Percentages		
White	0.8136	0.7718
Bachelor's degree	0.2209	0.1633
Southern residence	0.3456	0.3498
Non-Catholic Christian[a]	0.5264	0.6267
Catholic[a]	0.2564	0.2384
Jewish[a]	0.0182	0.0173
Other religion[a]	0.0309	0.0267
Means (standard deviations)		
Age	44.1831	48.4360
	(18.1325)	(19.4653)

[a] The religion reference category is no religion

The GOPS estimation found limited significant relationships between marital happiness and the spousal comparison variables. Having the same religion as one's spouse was positively associated with marital happiness (but significantly so for women only). For the same-educational-level variable, the signs of three of the four coefficients were positive as expected (but not significantly so); for the fourth coefficient, the sign was statistically significant at the ten percent level but of the wrong sign. Previous studies examining the relations between marital quality and spousal similarity have focused primarily on similarities in personality characteristics rather than similarities in demographic characteristics. Gattis et al. (2004) did, however, explore the correlation between spouses' education for distressed (treatment-seeking) as well as for non-distressed couples. They found no statistically significant differences in the education correlations for the two couple types. Gaunt's (2006) study included, not the spouses' particular religions, but a self-rating of religiosity, which was specified as: 1 = secular, 2 = traditional, 3 = orthodox, and 4 = ultra-orthodox. When the absolute value of the difference between the husband's and the wife's religiosity was included as an explanatory variable in a regression equation, Gaunt found marital satisfaction to be negatively related to the religiosity difference, but significantly so only for wives.

None of the coefficients of the socioeconomic index gap in the current study were found to be statistically significant. If the SEI gap reflects the extent to which household roles are traditional, we may gain insights by viewing other researchers' findings concerning related variables. Here we find mixed results. In their study, Tsang et al. (2003) included a traditionalism variable based on the proportion of household chores done by each partner. In their regression analysis, they found a significant effect of that explanatory

Table 6 GOPS Estimation results for marital status equation

Variable	Males function $Z\gamma$		Females function $Z\gamma$	
	Coefficient	p value	Coefficient	p value
Constant	−4.4161***	<0.0001	−3.9042***	<0.0001
Age	0.1348***	<0.0001	0.1199***	<0.0001
Age squared	−0.0011***	<0.0001	−0.0011***	<0.0001
White	0.3432***	<0.0001	0.5267***	<0.0001
Bachelor's degree	0.0746	0.1374	0.1671**	0.0017
Southern residence	−0.0482	0.2676	−0.0190	0.6597
Non-Catholic Christian	0.4280***	<0.0001	0.3167***	<0.0001
Catholic	0.3435***	<0.0001	0.2326**	0.0030
Jewish	0.3315*	0.0367	0.2441	0.1397
Other religion	0.2867*	0.0226	0.3944**	0.0031
Sample size	4500		4500	
Percent married/not married correctly predicted	65.7		67.4	

A positive coefficient in function $Z\gamma$ indicates that an increase in the specified variable increases the probability of being married

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 7 Conventional ordered probit estimation results for marital happiness

Variable	Function Xβ	
	Coefficient	p value
Males ($n = 1478$)		
Intercept 3	0.4437	0.2267
Intercept 2	2.1930***	<0.0001
Age	−0.0210	0.1291
Age squared	0.0002	0.1710
White	0.3937***	<0.0001
Number of children in household	−0.0456	0.1541
Lower class	−0.0752	0.6333
Upper class	0.4202†	0.0601
Hours worked	0.0000	0.9903
Spouse's hours worked	0.0000	0.9940
Same educational level	0.0271	0.6827
Same religion group	0.1922*	0.0206
Socioeconomic index gap	0.0020	0.2245
Females ($n = 1169$)		
Intercept 3	1.2819**	0.0015
Intercept 2	2.9012***	<0.0001
Age	−0.0462**	0.0025
Age squared	0.0004**	0.0080
White	0.2220*	0.0316
Number of children in household	−0.1595***	<0.0001
Lower class	−0.1599	0.3390
Upper class	0.0897	0.7377
Hours worked	0.0001	0.9545
Spouse's hours worked	−0.0005	0.8274
Same educational level	−0.0814	0.2674
Same religion group	0.2156 *	0.0185
Socioeconomic index gap	−0.0026	0.1406

A positive coefficient in function $X\beta$ indicates that an increase in the specified variable increases the probability of being very happy versus pretty happy or pretty happy versus not too happy

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 8 Generalized ordered probit (without selectivity) results for marital happiness

Variable	Function $X\beta_1$		Function $X\beta_2$	
	Coefficient	p value	Coefficient	p value
Males ($n = 1478$)				
Constant	2.2983*	0.0251	0.3344	0.3726
Age	−0.0316	0.4048	−0.0160	0.2585
Age squared	0.0003	0.4556	0.0001	0.3155
White	0.4858*	0.0122	0.3937***	<0.0001
Number of children in household	−0.0082	0.9179	−0.0483	0.1463
Lower class	−0.3064	0.3094	−0.0257	0.8761
Upper class	−0.1174	0.8030	0.4644*	0.0447
Hours worked	0.0136*	0.0296	−0.0013	0.5525
Spouse's hours worked	−0.0099	0.1385	0.0009	0.7200
Same educational level	0.0727	0.6582	0.0230	0.7366
Same religion group	0.1721	0.3813	0.1899*	0.0274
Socioeconomic index gap	−0.0045	0.3244	0.0025	0.1472
Females ($n = 1169$)				
Constant	2.4929**	0.0029	1.3258**	0.0016
Age	−0.0393	0.2386	−0.0461**	0.0035
Age squared	0.0004	0.2289	0.0004*	0.0133
White	0.1826	0.3443	0.2402*	0.0289
Number of children in household	−0.1761**	0.0031	−0.1590***	<0.0001
Lower class	−0.4487	0.1019	−0.0864	0.6245
Upper class	−0.3409	0.4674	0.1192	0.6667
Hours worked	−0.0032	0.5368	0.0007	0.7884
Spouse's hours worked	0.0055	0.2255	−0.0016	0.4953
Same educational level	0.2100	0.1618	−0.1314†	0.0882
Same religion group	0.0917	0.6121	0.2423*	0.0123
Socioeconomic index gap	−0.0035	0.3240	−0.0024	0.1915

A positive coefficient in function $X\beta_1$ indicates that an increase in the specified variable increases the probability of being pretty happy versus not too happy. A positive coefficient in function $X\beta_2$ indicates that an increase in the specified variable increases the probability of being very happy versus pretty happy

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

variable in only one of their three time periods; for that period, the effect on marital happiness was negative. Their partial path analysis showed no significant impact of that variable. In their full path model, they included a sex role traditionalism attitude variable that measured the extent of "belief in household roles for women and external work roles for men…." That variable had indirect effects on marital happiness; sex role traditionalism led to lower family income, which led to greater dissatisfaction with financial situation, which led to lower marital happiness. The work of Van den Troost et al. (2006) also included a sex role attitude variable that measured the degree of traditionalism. Their regression analysis found that the marital satisfaction of a woman was not significantly related to her own role traditionalism but was negatively related to her husband's. The marital satisfaction of a man was significantly related to neither his own role traditionalism nor to his wife's.

Without the GOPS procedure, it would not be possible to determine whether the relation between marital satisfaction and the explanatory variables differ with the level of marital satisfaction. A particular variable may be strongly related to an individual being pretty happy rather than not too happy, but only weakly or not at all related to an individual being very happy rather than pretty happy. F tests with the GOPS procedure enable us to determine whether these differences are statistically significant. Among men, significant differences were found (at the 10 % level or better), for hours worked and spouse's hours worked. Among women, significant differences were found for lower class, spouse's hours worked, and same education as spouse.[7]

The discussion above indicated whether the results of the GOPS analysis were consistent with those of other studies. However, given the differences in the variables employed as well as differences in the datasets, it is not possible to see how the GOPS methodology influences the findings. The following subsection directly compares the results of the GOPS procedure to those of other probit techniques.

Comparison of the Results from the GOPS Procedure and Other Probit Methods

The results of the conventional ordered probit are shown in Table 7, and the results of generalized ordered probit without selectivity are shown in Table 8. Recall that conventional ordered probit assumes that the effects of an explanatory variable are the same regardless of the satisfaction level. So while the coefficient results of generalized ordered probit (with or without selectivity correction) indicate separately (1) how a variable affects whether an individual is in satisfaction level 2 rather than satisfaction level 1, and (2) how a variable affects whether an individual is in satisfaction level 3 rather than satisfaction level 2, the corresponding results for conventional ordered probit indicate how a variable affects whether an individual is in either satisfaction level 2 or 3 versus the satisfaction level below that. When the results of GOPS and conventional ordered probit are compared (Table 7), we see that, for men, unlike the GOPS procedure, conventional ordered probit failed to find any significant association of marital happiness with age or age-squared, number of children, or hours worked by the respondent. For women, conventional probit estimation also failed to find any significant association of marital happiness with class. Also, unlike the GOPS procedure, conventional ordered probit found a statistically significant effect of same religion group for men and of white race for women.

When the results of GOPS and generalized ordered probit without selectivity are compared (Table 8), we see that the signs of all coefficients are the same for GOPS and generalized ordered probit without selectivity, with the exception of the coefficients of white in the women's equation. Focusing on significance at the 10 % level or better, however, there are several differences.

For men, the coefficients of age and age-squared, in both the $X\beta_1$ and $X\beta_2$ functions, and the number of children in the $X\beta_2$ function failed to be significant when selectivity is ignored, while they achieved significance when selectivity is taken into consideration. Looking more closely, we see that the magnitudes of those coefficients were larger when selectivity is taken into consideration than when it is ignored; so failing to adjust for selectivity has led to underestimating the size of the effects of those variables on the marital satisfaction of men. The coefficients of white and same religion group appeared to be significant in the $X\beta_2$ function when selectivity was ignored but not when selectivity was taken into consideration. Examining the coefficients shows that the effects of white and same religion group are overestimated, when selectivity is ignored.

For women, the coefficients of age, age-squared, and lower class in the $X\beta_1$ function failed to be significant when selectivity is ignored, while they achieved significance when selectivity is taken into consideration. The magnitudes of the coefficients of the age variables are smaller when selectivity is ignored than when it is taken into consideration. Thus, ignoring selectivity results in underestimating the effects of age on women's marital satisfaction. The magnitude of the coefficient of the lower class variable in the $X\beta_1$ function is actually larger when selectivity is ignored than when it is taken into consideration, indicating that ignoring selectivity results in overestimating its effect.[8] Also for women, the coefficient of white appeared to be significant in the $X\beta_2$ function when selectivity was ignored but not when selectivity was taken into consideration. The effect of white race appears to be overestimated when selectivity is ignored.

Summary and Conclusions

This paper developed the method of GOPS. While prior to the development of this procedure, it was possible to estimate the effects of explanatory variables on dependent variables such as marital satisfaction, it was necessary that the measure of marital satisfaction be a continuous variable, in order to take

[7] It should be noted that coefficients may be significantly different from each other, even though neither is significantly different from zero.

[8] This outcome occurs because significance not only depends on the magnitude of the coefficient estimate, but is also inversely related to the magnitude of the standard error estimate. In this case, the standard error is larger when selectivity is ignored, more than offsetting the somewhat larger coefficient estimate. So while the coefficient fails to show significance, the effect of the variable is actually overestimated.

sample selectivity into consideration. Heckman's (1976) method could then be used to make the needed adjustment. If the dependent variable was discrete, however, no appropriate technique was available. This procedure adds to our toolbox a methodology which enables researchers to handle situations in which the dependent variable is a discrete, ordered one and selectivity may be a problem. If researchers used conventional ordered probit to estimate the effects of a set of explanatory variables on an individual's discrete rating of marital satisfaction, they would be ignoring the potential sample selectivity issue as well as assuming equality of slopes; both of those actions can result in incorrect estimates of the relations of the explanatory variables to marital satisfaction.

Using data from the US General Social Survey, the GOPS method was demonstrated for estimating the relations between marital happiness and a set of explanatory variables. The GOPS results were compared to those from conventional ordered probit and generalized ordered probit without selectivity. Percentages correctly predicted were comparable for all three techniques. However, the conventional technique failed to find significance of some relationships identified by the GOPS method. Generalized ordered probit without selectivity also failed to find significant relationships for some variables. For other variables, some relationships were identified as statistically significant when selectivity was ignored but not when selectivity was taken into consideration. By comparing the findings of the different methods for the same data, it is shown that the conclusions drawn are affected by whether the researcher takes selectivity into consideration.

While other studies have found relationships between marital satisfaction and explanatory variables, without the GOPS procedure, it is not possible to determine whether the relations between marital satisfaction and the explanatory variables differ with the level of marital satisfaction. The results of this study suggest that these differences do exist. However, given the relatively small number of observations in the lowest satisfaction category, more research needs to be undertaken to corroborate the findings.

The paper showed the derivation of the relevant probability functions, from which the likelihood function is created. The SAS procedure NLMIXED can be used to perform the maximum likelihood estimation. Because the researcher needs to write several statements to set up the probabilities and the likelihood function, it is somewhat more difficult to do GOPS than to perform estimation using an established but potentially inappropriate procedure such as conventional ordered probit. Furthermore, depending on the sample and the specification, the second-order optimality condition may be violated and the researcher may need to explore a different specification.

It should be noted that the statistical results of any study are subject to random error and the conclusions drawn always depend on the sample. However, in general, the problem of selectivity implies that biases exist and making selectivity adjustments can be expected to lead to results which are closer to the true but unknown actual effects.

The technique presented in this paper can be applied not only to the analysis of marital satisfaction, but also to virtually any case in which respondents provide a discrete rating of a situation in which they have been involved. These cases include, but are not limited to, ratings of a job, a course, or a service. Thus, this procedure opens the door to analyses in numerous contexts.

Appendix

Letting ϕ denote a bivariate normal density function, Φ_1 a univariate cumulative normal distribution, and Φ_2 a bivariate cumulative normal distribution, the calculations of the probabilities are shown below.

$$
\begin{aligned}
\text{Case 1}: &\ \Pr(M = 1 \text{ and } S = 1) \\
&= \Pr(u > -Z\gamma \text{ and } \varepsilon \le -X\beta_1) \\
&= \int_{-Z\gamma}^{\infty} \int_{-\infty}^{-X\beta_1} \phi(u, \varepsilon)\, d\varepsilon \text{ du} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{-X\beta_1} \phi(u, \varepsilon)\, d\varepsilon \text{ du} \\
&\quad - \int_{-\infty}^{-Z\gamma} \int_{-\infty}^{-X\beta_1} \phi(u, \varepsilon)\, d\varepsilon \text{ du} \\
&= \Phi_1(-X\beta_1) - \Phi_2(-Z\gamma, -X\beta_1, \rho)
\end{aligned}
$$

$$
\begin{aligned}
&\text{Case } j \text{ (where } j = 2 \text{ to } s - 1) : \\
&\Pr(M = 1 \text{ and } S = j) = \Pr(u > -Z\gamma \text{ and } -X\beta_{j-1} \le \varepsilon \le -X\beta_j) \\
&= \int_{-Z\gamma}^{\infty} \int_{-X\beta_{j-1}}^{-X\beta_j} \phi(u, \varepsilon)\, d\varepsilon\, du \\
&= \int_{-\infty}^{\infty} \int_{-X\beta_{j-1}}^{-X\beta_j} \phi(u, \varepsilon)\, d\varepsilon\, du - \int_{-\infty}^{-Z\gamma} \int_{-X\beta_{j-1}}^{-X\beta_j} \phi(u, \varepsilon)\, d\varepsilon\, du \\
&= \left[\int_{-\infty}^{\infty} \int_{-\infty}^{-X\beta_j} \phi(u, \varepsilon)\, d\varepsilon\, du - \int_{-\infty}^{\infty} \int_{-\infty}^{-X\beta_{j-1}} \phi(u, \varepsilon)\, d\varepsilon\, du \right] \\
&\quad - \left[\int_{-\infty}^{-Z\gamma} \int_{-\infty}^{-X\beta_j} \phi(u, \varepsilon)\, d\varepsilon\, du - \int_{-\infty}^{-Z\gamma} \int_{-\infty}^{-X\beta_{j-1}} \phi(u, \varepsilon)\, d\varepsilon\, du \right] \\
&= \left[\Phi_1(-X\beta_j) - \Phi_1(-X\beta_{j-1}) \right] - \left[\Phi_2(-Z\gamma, -X\beta_j, \rho) \right. \\
&\quad \left. - \Phi_2(-Z\gamma, -X\beta_{j-1}, \rho) \right]
\end{aligned}
$$

Case s: $\Pr(M = 1 \text{ and } S$

$$= \Pr(u > -Z\gamma \text{ and } \varepsilon \geq -X\beta_{s-1})$$

$$= \int_{-Z\gamma}^{\infty} \int_{-X\beta_{s-1}}^{\infty} \phi(u,\varepsilon) d\varepsilon\, du$$

$$= \int_{-\infty}^{\infty} \int_{-X\beta_{s-1}}^{\infty} \phi(u,\varepsilon) d\varepsilon\, du - \int_{-\infty}^{-Z\gamma} \int_{-X\beta_{s-1}}^{\infty} \phi(u,\varepsilon) d\varepsilon\, du$$

$$= \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(u,\varepsilon) d\varepsilon\, du - \int_{-\infty}^{\infty} \int_{-\infty}^{-X\beta_{s-1}} \phi(u,\varepsilon)\, d\varepsilon\, du \right]$$

$$- \left[\int_{-\infty}^{-Z\gamma} \int_{-\infty}^{\infty} \phi(u,\varepsilon) d\varepsilon\, du - \int_{-\infty}^{-Z\gamma} \int_{-\infty}^{-X\beta_{s-1}} \phi(u,\varepsilon)\, d\varepsilon\, du \right]$$

$$= [1 - \Phi_1(-X\beta_{s-1})] - [\Phi_1(-Z\gamma) - \Phi_2(-Z\gamma, -X\beta_{s-1}, \rho)]$$

$$= 1 - \Phi_1(-Z\gamma) - \Phi_1(-X\beta_{s-1}) + \Phi_2(-Z\gamma, -X\beta_{s-1}, \rho)$$

$$\textit{Case}\; s+1: \Pr(M = 0) = \Pr(u \leq -Z\gamma) = \Phi_1(-Z\gamma)$$

References

Gattis, K. S., Berns, S., Simpson, L. E., & Christensen, A. (2004). Birds of a feather or strange birds? Ties among personality dimensions, similarity, and marital quality. *Journal of Family Psychology, 18*, 564–574. doi:10.1037/0893-3200.18.4.564.

Gaunt, R. (2006). Couple similarity and marital satisfaction: Are similar spouses happier? *Journal of Personality, 74*, 1401–1420. doi:10.1111/j.1467-6494.2006.00414.x.

Greene, W. H., & Hensher, D. A. (2008). *Modeling ordered choices: A primer and recent developments*. Cambridge: Cambridge University Press. doi:10.2139/ssrn.1213093.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5,* 475–492. Retrieved 9 Feb 2014 from http://www.nber.org/chapters/c10491 .

Maume, D. J., & Sebastian, R. A. (2012). Gender, nonstandard work schedules, and marital quality. *Journal of Family and Economic Issues, 33*, 477–490. doi:10.1007/s10834-012-9308-1.

Nakao, K., & Treas, J. (1992). *The 1989 Socioeconomic Index of Occupations: Construction from the 1989 Occupational Prestige Scores*. GSS Methodological Report No. 74. Retrieved 9 Feb 2014 from http://publicdata.norc.org:41000/gss/DOCUMENTS/REPORTS/Methodological_Reports/MR074.pdf .

Pedersen, D. E., & Minnotte, K. L. (2012). Dual earner husbands and wives: Marital satisfaction and the workplace culture of each spouse. *Journal of Family and Economic Issues, 33*, 272–282. doi:10.1007/s10834-012-9294-3.

Smith, T.W., Marsden, P.V., & Hout, M. (2011). General Social Survey, 1972-2010 [Cumulative File]. ICPSR31521-v1. Storrs, CT: Roper Center for Public Opinion Research, University of Connecticut/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors], 2011-08-05. doi:10.3886/ICPSR31521.v1 .

Tsang, L. L. W., Harvey, C. D. J., Duncan, K. A., & Sommer, R. (2003). The effects of children, dual earner status, sex role traditionalism, and marital structure on marital happiness over time. *Journal of Family and Economic Issues, 24*, 5–26. doi:10.1023/A:1022478919443.

Van den Troost, A., Matthijs, K., Vermulst, A. A., Gerris, J. R. M., & Welkenhuysen-Gybels, J. (2006). Effects of spousal economic and cultural factors on Dutch marital satisfaction. *Journal of Family and Economic Issues, 27*, 235–262. doi:10.1007/s10834-006-9016-9.

Karen Leppel is a Professor of Economics at Widener University. She completed her Ph.D. in Economics at Princeton University. Her research interests pertain to the impact on economic behavior of demographic characteristics such as gender, race, ethnicity, age, and sexual orientation. She was the recipient of the 2011–2012 Widener University School of Business Administration Distinguished Research Professor Award. She also serves on the Board of Editors of International Advances in Economic Research.