The Detection and Correction of Sample Selection Bias
in Patient Satisfaction Surveys:
A Two-Step Application of PROC PROBIT

Guy Pascale, Senior Project Manager, Keystone Mercy Health Plan, Philadelphia, PA
Howard Peters, Research Analyst, Virtua Health-Memorial Hospital, Mount Holly, NJ

Abstract:

With increased emphasis on quality in the hospital setting, the use of patient satisfaction surveys has increased dramatically over the last decade. Although these survey instruments can identify problem service areas, little is known about the impact of patient non-response. The authors analyze the impact of non-response on an in-patient satisfaction survey for a community hospital by utilizing a two-stage process. First, the authors model response versus non-response against patient, payor, and hospital specific characteristics using PROC PROBIT. This information is used to calculate the inverse Mills ratio (IMR) for all observations. The second stage entails modeling respondent likelihood of recommending the hospital via an ordered PROC PROBIT. The IMR is included as an independent variable to test and correct for non-response bias. If the coefficient on the IMR is insignificant, then non-response bias does not exist. If the coefficient on the IMR is significant, non-response bias does exist. However, the inclusion of the IMR controls for the presence non-response bias. Thus, the model is properly specified. The authors include a programming example.

Introduction:

With the upheaval of the managed care revolution in health care, hospitals have sought better ways to monitor the quality and effectiveness of the services they provide. These have included benchmarking length of stay, cost, readmission rates, infection rates, and other quality indicators against a peer group or a peer hospital. Also of importance is the measurement of patient satisfaction. The analysis of patient satisfaction is an important component to quality assurance in the hospital as it provides insight into areas where patients feel the quality of their care is lacking. The medical literature contains much in terms of patient satisfaction. The below literature review is not exhaustive but serves to illustrate the variety of patient satisfaction research that appears and also to points to areas where more is research is necessary.

Much research has been done on survey methodology. Some authors have investigated the consistency of survey results across different survey instruments and found good agreement among surveys with different wording (9). Others have written on the development and use of patient satisfaction surveys (10,12,18) while other authors have analyzed the impact of sponsorship on patient satisfaction surveys (16). Some have stressed the use of patient satisfaction surveys as a marketing research tool for general practices (26).

Patient satisfaction with in-patient procedures is also an important
component of the literature. This area of research mainly addresses all in-patients (22). Others have investigated satisfaction with vitreoretinal surgery (35), total hip arthroplasty (29), and the informed consent process (1).

Satisfaction with out-patient procedures and emergency room patients also appears in the literature. Authors have investigated satisfaction with general out-patient services (2) but also includes specific medical conditions including eye surgery (5), cancer (39), HIV (27), and conscious sedation for bronchoscopy (33). The literature on emergency room satisfaction centers on factors that influence satisfaction (7,45).

Research into patient satisfaction also includes physician office practices (15,18,26,36,37,42), mental health (3,17,20,25,34,44), and nursing (10,34).

One critical flaw with all the literature listed above is the lack of discussion about the impact of response rates and non-response bias. Some authors have sought to address these issues.

Barkley et al has stated that "a low response rate on a patient satisfaction survey can yield strikingly different results from those obtained with a larger group of respondents, which critically affects the identity of the area targeted for improvement" (4). The authors come to this conclusion by comparing summary statistics of the first 30% of respondents to all respondents to their survey. They find that 9 out the 13 scales measured exhibited a significant change due to the increased sample size.

Coversely, Lasek et al has found a minimal impact of low response rates and non-response bias (28). The authors mailed patient satisfaction surveys to in-patients 8 to 12 weeks after discharge from 29 hospitals. If in 4 weeks a response was not received, a second questionnaire was mailed. If no response was received to the second surveys after 4 more weeks, telephone surveys were initiated. After analysis, the authors conclude "that impact of non-response bias on satisfaction surveys of hospitalized patients may be relatively small and that the effects were not systematically greater in hospitals with lower response rates".

Other authors have investigated this issue without discussing concrete statistical solutions to the sample selection issue (6,11,25,30,31,32,38,42).

Since the literature lacks a statistical approach to sample selection bias detection/correction, we intend to introduce an application of methodologies that addresses the bias issue with regards to patient satisfaction surveys.

**Methodology:**

The issue of sample selection bias remained relatively untouched until the publication of the Heckman article in 1979 (23). This seminal article addresses the sample selection issue by developing a two-stage process where stage one entails fitting a PROBIT, using all observations, to determine the probability of being a non-limit observation.

From the PROBIT, the inverse Mills ratio (IMR), or hazard rate, is calculated for all observations. The IMR can be considered as a measure of how much
information on the dependent variable is discarded. Dependent on the type of censoring, the IMR is specified as follows:

\[ \lambda_{\text{lo}} = \frac{f(\phi)}{1 - F(\phi)} \text{ or } \lambda_{\text{hi}} = \frac{f(\phi)}{F(\phi)} \]

where \( \lambda_{\text{lo}} \) is the IMR for censoring at the lower portion of the distribution, where \( \lambda_{\text{hi}} \) is the IMR for censoring at the upper portion of the distribution, where \( f(\phi) \) is the normal probability density function, and where \( F(\phi) \) is the normal cumulative density function.

The second stage is a standard OLS regression, on the non-limit observations, with the IMR utilized as an independent variable. The inclusion of the IMR controls for the characteristics that discriminate between response and non-response. This corrects for any non-response bias.

This methodology has been utilized in the medical literature to adjust for non-response bias in a survey of medicine use by the elderly (21).

This methodology is fine if you have a continuous measure as dependent variable (i.e. cost, earnings, etc.). But when the dependent measure is discrete (as in our case), some modifications must be made. This issue is addressed by Eklof and Karlsson (14). The authors address non-response in discrete choice models for contingent valuation studies. They use Monte-Carlo simulations to determine the power of Likelihood Ratio, Wald, LM, and an Omitted Variable (OV) test. The OV test is a modification of the Heckman method (40). The specification of the OV variable is as follows:

\[ OV = \frac{f(\phi)}{F(\phi)}. \]

Note that this is the same as the IMR for censoring for the upper part of the distribution.

The authors find "that the computationally straightforward OV test, which does not require maximization of the full bivariate probit likelihood, performs well. It has good size properties and the size adjusted power is comparable to the power of the other tests". Thus the modified Heckman technique is a viable technique for sample selection bias detection/correction for patient satisfaction surveys.

**Data:**

Virtua Health-Memorial Hospital of Burlington County is a 350 bed facility and is part of a five hospital system. The system also includes a skilled nursing facility (two as of September 1999) and two out-patient surgical centers.

The hospital uses the Press-Ganey survey for patient satisfaction. This is a standardized survey that is utilized by approximately 1000 hospitals nationwide. The surveys are mailed to all in-patients (excluding newborns, mental health patients, and the deceased) approximately two weeks after discharge. The patient's ID number is included on the survey instrument. This allows the linkage of survey information to all hospitalization data. If a survey is not returned, the survey data is listed as missing.
Stage 1: Response/Non-Response Analysis:

Appendix 1 contains the PROC PROBIT analysis of responders versus non-responders.

Many factors influence the probability of response to the survey. Longer length of stay or greater severity of illness are associated with a lower probability of response. Medicaid, self-pay, and charity care patients also have a lower probability of response. Patients in DRG's 127, 383, 98, and 210 have a lower probability of response (CHF, Other Antepartum DX, Bronch/Asthm Age <17, and Hip/Femur Age > 17, respectively).

Patients who are married or white have a higher probability of response. Patients in DRG's 148, 122, 359, and 15 also have a higher probability of response (Small/Large Bowel Procedures, Circulatory Disorders, Non-Malignant Uterine Procedures, and Transischemic Attacks, respectively).

Stage 2: Sample Selection Detection/Correction:

Appendix 2 contains the PROC PROBIT analysis of likelihood of recommending the hospital.

Married patients and females who gave birth give the hospital higher recommend scores that other patients. The IMR is not significant in predicting the recommend score, therefore, sample selection bias does not exist.

Conclusions:

The application of the modified Heckman (23) technique posited by Vella (40) is a straightforward correction for sample selection bias in discrete choice models. This allows for the detection and correction of bias in patient satisfaction surveys. The application of this technique could save health care organizations money in terms re-surveying non-respondents.
Appendix 1: PROC PROBIT analysis of responders versus non-responders

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Chi-Square</th>
<th>Pr&gt;Chi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.170</td>
<td>1740.56</td>
<td>0.0001</td>
</tr>
<tr>
<td>AgeLT18</td>
<td>0.034</td>
<td>0.56</td>
<td>0.4514</td>
</tr>
<tr>
<td>DaysGt6</td>
<td>-0.137</td>
<td>29.39</td>
<td>0.0001</td>
</tr>
<tr>
<td>S34</td>
<td>-0.179</td>
<td>64.72</td>
<td>0.0001</td>
</tr>
<tr>
<td>Married</td>
<td>0.325</td>
<td>323.66</td>
<td>0.0001</td>
</tr>
<tr>
<td>White</td>
<td>0.311</td>
<td>146.04</td>
<td>0.0001</td>
</tr>
<tr>
<td>Medicaid</td>
<td>-0.452</td>
<td>53.03</td>
<td>0.0001</td>
</tr>
<tr>
<td>Self-Pay</td>
<td>-0.757</td>
<td>78.28</td>
<td>0.0001</td>
</tr>
<tr>
<td>Charity</td>
<td>-0.239</td>
<td>12.74</td>
<td>0.0004</td>
</tr>
<tr>
<td>Military</td>
<td>-0.286</td>
<td>7.41</td>
<td>0.0065</td>
</tr>
<tr>
<td>Mother</td>
<td>0.102</td>
<td>16.21</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG 127</td>
<td>-0.081</td>
<td>3.50</td>
<td>0.0611</td>
</tr>
<tr>
<td>DRG 209</td>
<td>0.355</td>
<td>41.82</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG 182</td>
<td>-0.113</td>
<td>2.95</td>
<td>0.0858</td>
</tr>
<tr>
<td>DRG 138</td>
<td>0.088</td>
<td>1.66</td>
<td>0.1966</td>
</tr>
<tr>
<td>DRG 148</td>
<td>0.618</td>
<td>74.68</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG 383</td>
<td>-0.571</td>
<td>33.19</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG 122</td>
<td>0.243</td>
<td>10.12</td>
<td>0.0015</td>
</tr>
<tr>
<td>DRG 98</td>
<td>-0.254</td>
<td>5.39</td>
<td>0.0202</td>
</tr>
<tr>
<td>DRG 359</td>
<td>0.448</td>
<td>30.91</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG 15</td>
<td>0.245</td>
<td>8.38</td>
<td>0.0038</td>
</tr>
<tr>
<td>DRG 141</td>
<td>0.143</td>
<td>2.53</td>
<td>0.1112</td>
</tr>
<tr>
<td>DRG 494</td>
<td>0.128</td>
<td>1.99</td>
<td>0.1578</td>
</tr>
<tr>
<td>DRG 210</td>
<td>-0.255</td>
<td>5.09</td>
<td>0.0240</td>
</tr>
</tbody>
</table>
Appendix 2: PROC PROBIT analysis of likelihood to recommend

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Chi-Square</th>
<th>Pr&gt;Chi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.3004</td>
<td>22.85</td>
<td>0.0001</td>
</tr>
<tr>
<td>Married</td>
<td>0.1056</td>
<td>6.01</td>
<td>0.0142</td>
</tr>
<tr>
<td>Medicaid</td>
<td>-0.1960</td>
<td>1.88</td>
<td>0.1694</td>
</tr>
<tr>
<td>Mother</td>
<td>0.2214</td>
<td>23.34</td>
<td>0.0001</td>
</tr>
<tr>
<td>DRG182</td>
<td>-0.2153</td>
<td>3.11</td>
<td>0.0777</td>
</tr>
<tr>
<td>DRG383</td>
<td>-0.2746</td>
<td>1.64</td>
<td>0.1995</td>
</tr>
<tr>
<td>DRG122</td>
<td>0.2354</td>
<td>3.17</td>
<td>0.0748</td>
</tr>
<tr>
<td>DRG141</td>
<td>0.2827</td>
<td>2.83</td>
<td>0.0922</td>
</tr>
<tr>
<td>DRG97</td>
<td>0.2911</td>
<td>1.77</td>
<td>0.1828</td>
</tr>
<tr>
<td>IMR</td>
<td>-0.0083</td>
<td>0.00</td>
<td>0.9622</td>
</tr>
</tbody>
</table>
Appendix 3: Two Stage Probit Program

libname guy 'd:\sasdata';

data one;
  set guy.nesug2;
  if sever gt 2 then s34 = 1; else s34 = 0;
  if respond eq 1 then nrespond = 0; else nrespond = 1;

proc probit;
  class nrespond;
  model nrespond = agelt18 daysgt6 s34 married white
    medicaid selfpay charity military
    mother drg127 drg209 drg182 drg138 drg148 drg383
drg122 drg98 drg359 drg15 drg141 drg494 drg210
    / converge = .00001;
  output out = probit2 xbeta = xbpr prob = probpr;

data probit2;
  set probit2;
  pdfnorm = exp (-.5*xbpr*xbpr)/sqrt(2*3.1459);
  millslo = pdfnorm/(1-probnorm(-xbpr));
  millshi = pdfnorm/probnorm(-xbpr);

proc means;
  var pdfnorm millslo millshi;

data probit3;
  set probit2;
  if recomm ne 0;
    if recomm eq 5 then nrecomm = 1;
    if recomm eq 4 then nrecomm = 2;
    if recomm eq 3 then nrecomm = 3;
    if recomm eq 2 then nrecomm = 4;
    if recomm eq 1 then nrecomm = 5;

proc probit;
  class nrecomm;
  model nrecomm = married
    mother drg182 drg122 drg141 millshi
    / converge=.00001;

run;
Works Cited:

21) Grotzinger, K etal, "Assessment and Control of Non-response Bias in a Survey of Medicine Use by the Elderly", Medical Care, 1994, 32 (10), pages 989-1003.
26) Khayat, K, "Patient Satisfaction Surveys as a Market Research Tool for General Practices