PROBABILITY OF RESPONSE: A MULTIVARIATE PROBIT ANALYSIS*

Sara L. Thran, William D. Marder, and Richard J. Willke, American Medical Association

## Introduction

Responsible use of most survey data depends on a clear understanding of the relationships between the group of survey respondents and the target population. Complicating the analysis of this relationship is the fact that, even within a survey, the group of respondents will differ from item to item. Thus, response analysis and associated weighting of the data collected in sample surveys has a long history. This study utilizes econometric techniques to develop an alternative weighting methodology. These methods are applied to the American Medical Association's telephone surveys on socioeconomic characteristics of medical practice.

The AMA Physician Masterfile, which is a continuously updated enumeration of all U.S. physicians, provides us with a unique opportunity to compare nonrespondents and respondents to our survey. Because we have access to Masterfile records for all nonrespondents we can analyze response patterns based on differences in key demographic and practice characteristics that are maintained on the file.

Most studies of nonresponse that attempt to correct the problem via appropriate weighting of survey data are forced to assume that nonrespondents are identical to respondents within strata. This assertion flies in the face of the known difference in their willingness to respond respond. Recent developments in econometrics can be applied to this problem. Heckman (1976) has developed regression techniques that help correct for non-random self selection. We apply these techniques to the self selection embodied in survey nonresponse.

We investigate the probability of responding to SMS surveys using multivariate probit estimation. Given our extensive knowledge of the characteristics of nonrespondents, we are able to measure key determinants of the probability of response. In addition to investigating survey nonresponse, we use the multivariate probit technique to investigate the willingness to respond to two key items on the SMS survey, income and hours worked. Following Heckman we use the estimated underlying and unobservable prediliction to respond in the construction of an instrument that we use to estimate unbiased coefficients of the determinants of income and hours worked. We use these coefficients to predict responses for nonrespondents and find relatively small corrections occur when our estimates are compared to an ordinary least squares weighting strategy.

## Previous Research on Nonresponse

Previous studies have compared demographic characteristics of respondents and nonrespondents. In most cases, information about individuals who do not respond to surveys is limited. Various strategies have been used to estimate characteristics of nonrespondents. Fitzgerald and Fuller (1982) report that interviewers gathered information on the designated respondent even if he refused to be interviewed. In cases where the designated respondent terminated the initial contact before supplying all the information, interviewer observation was used for some of the items. Differences were found in demographic characteristics of respondents and refusers -- refusers differed from respondents in terms of urbanization of the community, age, marital status, and dwelling type.

Smith (1983) examined several methods to estimate characteristics of nonrespondents in the 1980 General Social Survey. These methods included interviewer estimates and extrapolation for difficulty. According to interviewer estimates, nonrespondent households had older heads of household, fewer adults, and higher family incomes. In extrapolating for difficulty, it was found that labor force participation and high social status led to more difficulty; the young and urban were also more difficult to interview. Each method examined proved to be of limited usefulness in estimating nonrespondent characteristics.

Berk (1985) compared early and late respondents to the Physicians' Practice Survey, a component of the National Medical Care Expenditure Survey. Adding late respondents did not substantially affect most of the estimates of key demographic variables, leading Berk to conclude that making many callbacks over an extended period did not reduce nonresponse bias.

Item nonresponse has received little attention in the literature. Ferber (1966) analyzed item nonresponse by characteristics of respondents in a consumer mail survey and found that the pattern of nonresponse was very similar on different types of questions. Bell (1984) examined item nonresponse to income questions in a telephone survey. Logistic multiple regression showed that females, older respondents, and respondents with less education had a lower probability of answering the continous-scale income question.

Lillard et al. (1986) examined income item nonresponse in the 1980 Current Population Survey (CPS) and the Census Bureau imputation procedures that attempt to address the issue. They utilized multivariate probit estimation of the probability of responding to the income question on the CPS. As regressors in the probit equation they chose variables such as education and experience categories that are known to be determinants of the level of income. They concluded that Census imputation procedures, which ignore these issues, severely understate income in certain occupations and most-likely understate average income as well.

Unlike most other research on nonresponse, this study has information on the demographic characteristics of nonrespondents. Olson, et al. (1986) took the initial step of analyzing differences between SMS survey respondents and nonrespondents in a univariate context. This paper takes the further steps of reporting results of multivariate analyses of survey and item nonresponse and estimating nonresponse bias with regression techniques.

Description of the Socioeconomic Monitoring System (SMS) Program

The SMS program is a series of telephone surveys of physicians begun in 1981. SMS periodically collects information regarding medical practice characteristics Several SMS surveys are conducted each year including an annual core survey in the spring, which collects data from approximately 4,000 physicians through a 25-minute interview.

The sample for each survey is selected from the AMA Physician Masterfile. The eligible sample is limited to nonfederal patient care physicians, excluding resident physicians. The sample design is a stratified random sample, with the strata defined by specialty and geographic region. Each survey includes reinterviews with physicians who were initially interviewed a year earlier, as well as interviews with physicians who were selected for the first time.

This study examines 4340 eligible physicians who were sampled for the first time in the 1985 core survey. The survey was conducted by Mathematica Policy Research from March through June 1985. The response rate for the initial sample was 56.5%.

Analysis

We wish to examine the determinants of survey response and item response, and to test whether these determinants are different in the two situations. In contrast to previous studies, this analysis will be conducted in a multivariate context. We specify a "responsiveness" index $\beta_1 X_{1i} + \epsilon_{1i}$, where $\beta$ is a (row) vector of coefficients, $X_{1i}$ is a (column) vector of characteristics of individual i, and $\epsilon_{1i}$ is a normally distributed random error. It is defined such that the individual responds if the value of this index is above some critical value.[2] We consider survey response in conjunction with one item response at a time, and for exposition purposes we choose income item response. The survey and hours item response analysis will be exactly analagous.

Three sets of nested, successively less restricted hypotheses will be considered and tested against each other. These cases are: (1) the random multinomial, or sequential binomial model, (2) the single index, ordered response model, and (3) the double index, sequential response model.

The first, most restricted, case allows no individual characteristics to affect response probability, effectively restricting all slopes to be zero but allowing separate constant probabilities. It can be viewed as a multinomial model with three cases - no response at all, survey response but no income response, and income as well as survey response. This model could be a result of a structure in which all $\beta_1=0$, and the three types of response occur when (i) $\epsilon_{1i} < c_1$, (ii) $c_1 \leq \epsilon_{1i} < c_2$, and (iii) $c_2 \leq \epsilon_{1i}$, where $c_1$ and $c_2$ are constants.

If $\pi_1$, $\pi_2$ and $\pi_3$ are the resulting probabiliities of these three outcomes, respectively, and $N_1$, $N_2$ and $N_3$ are the number of sample individuals falling in each category, the log likelihood of the sample is $(N_1 \log\pi_1 + N_2 \log\pi_2 + N_3 \log\pi_3)$.

This case can also be viewed as a sequential binomial model, with survey response preceding income item response, the log likelihood of survey response is $(N_1 \log\pi_1 + (N_2+N_3)\log(\pi_2+\pi_3))$ and the log likelihood of item response, conditional on survey response, is $(N_2 \log(\pi_2/(\pi_2+\pi_3)) + N_3 \log(\pi_3/(\pi_2+\pi_3)))$. When combined, the log likelihoods reduce to the multinomial model. This log likelihood may be estimated using sample proportions.

The single index, ordered response model considers responsiveness to be a one-dimensional attribute, with the three categories representing low, moderate and high levels of responsiveness. The index parameters are scaled such that a person does not respond at all when $\beta_1 X_{1i} + \epsilon_{1i} < c_1$, the individual responds to the survey but not to the income item when $c_1 \leq \beta_1 X_{1i} + \epsilon_{1i}$, and responds to income as well when $c_2 \leq \beta_1 X_{1i} + \epsilon_{1i}$. This model thus allows the probability of a category to be determined by individual characteristics as well as a constant, but restricts the characteristics to affect survey and item response in the same way. The parameters $\beta_1$ and $c_2$ ($c_1$ is normalized to zero) can be estimated via an ordered probit maximum likelihood technique, and the resulting maximized log likelihood can be used to test the validity of this model versus the random multinomial model.

The double-index model corresponds structurally to the sequential binomial model. The index $\beta_1 X_{1i} + \epsilon_{1i}$ is relevant only for survey response: a person responds to the survey if and only if $\beta_1 X_{1i} + \epsilon_{1i} > c_1$, and does not respond when $\beta_1 X_{1i} + \epsilon_{1i} < c_1$. A new index $\beta_2 X_{2i} + \epsilon_{2i}$ represents responsiveness to the income question for those who respond to the survey. Thus, given $\beta_1 X_{1i} + \epsilon_{1i} > c_1$, a person responds to the income item if and only if $\beta_2 X_{2i} + \epsilon_{2i} > c_2$.

This structure allows slope coefficients to differ between the two types of response categories and essentially makes responsiveness a two-dimensional characteristic. It can be compared directly to the random sequential binomial model, but is also a less restricted version of the single-index model and can be compared to it as well.[3] This latter comparison proves interesting when deciding whether nonrespondents to the income item should be grouped together or separately by survey response or nonresponse. The survey and income item response parameter vectors can be estimated via a binomial probit on survey response for the entire sample, and a separate binomial probit for income item response using only survey respondents.[4] The sum of their maximized log likelihoods is then comparable to the log likelihoods of the other two models, and a likelihood ratio test may be performed.

The second phase of this analysis examines how item outcomes, namely income and hours, vary for nonrespondents vs. respondents. The fact that income, e.g., is known only for those who choose to respond to the survey and the income question, suggests these income responses may be subject to sample selection bias as described by Lillard et al. (1986). The problem arises because the fact that they choose to respond may be correlated with their incomes in a way not represented by observable characteristics. For example, "busy" doctors may have higher incomes and be less likely to take the time necessary to participate in the survey. To the extent that observable characteristics, or the random error in earnings, are correlated with the response probability, inconsistent estimates of the parameters in an income regression will be the result of ordinary least squares analysis, and simple weighting schemes for predicting nonrespondent income may not be appropriate. Estimates of the determinants of income can be derived with corrections for this bias following Heckman (1976).

The usual approach to this problem is to estimate a response probit or probits, as we have proposed, to construct a variable $\lambda$ which represents the expected value of the response error given the response decision (the inverse Mills' ratio), and include it in the income regression as a regressor.[5] Hence we construct $\lambda_1$, based on the survey response results, and $\lambda_2$ based on the income item response results and include both as regressors.[6] This procedure, by accounting for bias introduced by the response decisions, should result in consistent estimates of the parameters of the earnings function. Using these parameter estimates, we can predict income for nonrespondents.

Although it is tempting to construct $\lambda_1$ and $\lambda_2$ for nonrespondents to represent a responsiveness factor, and to use them to predict income, such a scheme is not feasible here, for two reasons. First, $\lambda_2$ is not defined for survey nonrespondents, since they are not included in the item response probit. Secondly, among those with income present, $\lambda_1$ and $\lambda_2$ are strictly positive by construction, resulting in coefficient estimates based solely on $\lambda$'s in that range. However, $\lambda_1$ must be negative for survey nonrespondents, and $\lambda_2$ must be negative for income nonrespondents. This presents a severe out-of-sample prediction problem.

Therefore, while we estimate the following regression:

$$Y = \alpha X_i + \sigma_1\lambda_1 + \sigma_2\lambda_2 + \mu_i$$

we predict income using only $\alpha X_i$, which are the true earnings parameters purged of selection bias. We then compare mean predicted incomes for the three categories of sample individuals. However, the exclusion of two regression variables (with non-zero means) from the prediction causes mean predicted income not to replicate mean actual income, so we compare using differences in predicted means only. A further comparison can be made using predicted incomes based on earnings parameters estimated without the selection bias corrections.

## Estimation Results

The demographic and practice characteristics examined are: specialty, census division, location, years since completion of undergraduate medical education, sex, board certification, AMA membership, country of medical school, and major professional activity. In addition, two direct measures of survey effort and responsiveness are included as explanatory variables in the probits, but are excluded from the income and hours regressions to provide structural identification of the $\lambda$ terms. First, the total number of calls made is used. Second, a response index is constructed for each item based on the number of other questions that the particular respondent answered.[7]

Of the 4340 physicians selected for the 1985 SMS core survey, 1889, or 43.5%, chose not torespond at all. Among the 2451 respondents, 629 did not respond to the income question and 107 did not answer the hours question. The sample log-likelihoods under the random multinomial or sequential binomial model are thus -4367.6 in the income case and -3411.5 in the hours case.

The results of the single-index ordered response probits are shown in Table 1. The first column of results are for the survey and income item response combination, while the second column represents the survey and hours item response combination. The constructed response index is not used here because it is not applicable for survey non-respondents.

The two sets of results are very similar, partly because they both include the survey response decision. Pediatricians, psychiatrists and anesthesiologists are the most likely to respond, while internists and surgeons are least likely. Rural and hospital-based physicians, FMGs, and AMA members are also more responsive. Physicians in western regions are more likely to respond than those in New England. The total number of calls made to the physician is negatively related to responsiveness. Although in some cases the physician gave a flat refusal on the first call, and no more calls were attempted, it apparently was much more common to make a number of calls before getting a refusal or giving up. Finally, responsiveness is a non-linear function of years since completion of undergraduate medical education, and declines until it bottoms out at about 28 years since M.D., after which responsiveness increases. The estimated critical level c is positive and very significant in both cases indicating in at least one dimension survey and item response are ordered manifestations of a general responsiveness attribute.

The $\chi^2$ statistics reported represent likelihood ratio tests of the ordered probit models vs. the random multinomial models, and in each case significantly reject the null hypothesis of zero slope coefficients.

Table 2 reports results of the binomial probit analyses on survey and item response. Pediatricians, radiologists, psychiatrists, and anesthesiologists are the most likely to respond to the survey while internists are the least likely. Rural physicians and physicians in the West South Central and Pacific divisions are more likely to respond. In addition, physicians with a large number of years since M.D., AMA members, FMGs, and hospital-based physicians are more likely to respond. The greater the number of calls made, the less likely is the physician to respond.

Among the survey respondents, the characteristics significantly related to the probability of responding to the annual income question are: specialty, census division, sex, country of medical school, total calls and, when used, the constructed response index variable. The variables related to response to the hours item are specialty, number of calls, and the constructed response index variable.

Specialty, census division, country of medical school, and number of calls made are related to both the probability of survey response and the probability of responding to the income question. However, although FMGs are more likely to respond to the survey, they are less likely to respond to the income item. Physicians in the West South Central division are more likely to respond to the survey and less likely to respond to the income question than physicians in New England. Response to the hours question cannot be predicted very accurately based on the characteristics studied here.

To test the single-index model versus the double-index model, we add the log-likelihoods of the survey and item response probits to provide a total log likelihood for the double-index model, and perform the likelihood ratio test. For the income case, this sum is -4133.2, against the ordered probit log-likelihood of -4177.6, resulting in a $\chi^2$ statistic of 88.9 with 27 degrees of freedom. Since the .01 critical value is 47.0, we can reject the single-index null hypothesis at that level. For the hours case, this sum is -3210.3, compared to the ordered probit log-likelihood of -3248.4. The $\chi^2$ statistic here is 76.3, again allowing rejection of the single-index hypothesis. Evidently the factors governing survey responsiveness are of a sufficiently different nature from those governing item responsiveness that they cannot be represented by a single index.

Table 3 presents the results of log net income and log total hours regressions including $\lambda_1$ and $\lambda_2$, the expected values of the responsiveness index errors, as regressors. The coefficient estimates of the observable characteristics are not of particular interest here, in that they present no surprises. The coefficient of $\lambda_1$ is essentially an estimated covariance of the earnings equation error with the survey response index error, while that of $\lambda_2$ is the covariance of the earnings equation error with the item response index error. The coefficient for $\lambda_1$ is statistically significant in both the income and the hours regressions. The result for income is similar to the results reported by Lillard et al. (1986). This would say that an income respondent who had a low income response index, and hence a more positive response error, would have higher income than otherwise predicted. This result is reasonable in that busy doctors may be less likely to respond, but if they do they reveal higher incomes.[8]

The results for $\lambda_1$ in the hours regression indicate that physicians with high survey response index errors - those who seem unlikely to respond but do - work longer hours. This makes sense in that the more hours they work, the more likely they are to be eventually caught by the interviewer.

The regression coefficients from Table 3 can be used to predict income and hours for all sample individuals. Although use of the selectivity correction coefficients for individual prediction presents difficulties described earlier, under the assumption of zero means for response errors across the entire sample, correct predicted means for this sample can be constructed using the other variable coefficients and means, with the $\lambda$ means equal to zero. Table 4 lists these results and compares them to the actual means for item respondents, and to the predicted means using the uncorrected regressions.

Using the sample standard error as a yardstick (the correct standard errors of the predicted means in the selectivity corrected regressions are quite complex to compute here), we find that use of selectivity corrections does not significantly affect predicted sample means. While the predicted sample means for income are more than one standard error above the item respondent mean, the two sample mean predictions are about one-half the standard error different. In the hours case, all three means are less than one-half the standard error apart. A priori one might expect the uncorrected regression to understate the earnings (or hours) effect of observable characteristics if respondents have downwardly biased earnings or hours, resulting in a lower predicted sample mean based on the uncorrected regression coefficients. Just the opposite is true of both variables, however, although the differences are small, allowing for no strong interpretation of these results.

## Conclusion

This study analyzed the determinants of survey and item non-response and the effects of nonresponse bias on net income and total hours practiced for a sample of physicians. Our tests indicate that for both variables the determinants of survey non-response are significantly different from the determinants of item non-response, and that the survey response decision is better explained by the independent variables than are the item response decisions. However, the corrections for response selection bias in income and hours regressions are of marginal significance, and do not result in significantly different predictions of population means. Hence the value of these corrections for this purpose, at least in the case where considerable information on non-respondents is available, appears quite limited.

*The authors gratefully acknowledge the comments and suggestions made by John Czajka of Mathematica Policy Research.

## References

Bell, Ralph (1984), "Item Nonresponse in Telephone Surveys: An Analysis of Who Fails to Report Income," Social Science Quarterly, 65(1), Mar, 207-215.

Berk, Marc (1985), "Interviewing Physicians: The Effect of Improved Response Rate," American Journal of Public Health, 75(11), Nov, 1338-1340.

Catsiapis, G. and Robinson, C. (1982), "Sample Selection Bias with Multiple Selection Rules: An Application to Student Aid Grants," Journal of Econometrics, 18(3), April.

Ferber, Robert (1966), "Item Nonresponse in a Consumer Survey," The Public Opinion Quarterly, 30(3), Fall, 399-415.

Fitzgerald, R. and Fuller, L. (1982), "I

Hear You Knocking but You Can't Come In: The Effects of Reluctant Respondents on Refusers and Sample Survey Estimates," Sociological Methods and Research, 11(1), Aug, 3-32.

Heckman, James (1976), "The Common Structure of Statistical Methods of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," Annals of Economic and Social Measurement, 5(4), Fall.

Lillard, Lee, Smith, J., and Welch, F. (1986), "What do we really know about wages? The importance on non-reporting and Census imputation," Journal of Political Economy, 94(3), June.

Olson, Lorayn, Thran, S., and Strouse, R. (1986), "Obtaining a Representative Data Set: The American Medical Association's Socioeconomic Monitoring System," Paper presented at annual meeting of the American Association for Public Opinion Research, St. Petersburg, May.

Smith, Tom (1983), "The Hidden 25 Percent: An Analysis of Nonresponse on the 1980 General Social Survey," The Public Opinion Quarterly, 47(3), Fall, 386-404.

## Footnotes

1. The Periodic Survey of Physicians (PSP), the AMA's precursor to SMS, was an annual mail survey of physicians which collected similar data. In an effort to shorten the field period as well as the time lag between the completion of the field period and the availability of results, the SMS telephone survey replaced PSP.
2. The deterministic part of this index is analagous to the discriminant vector of classical discriminant analysis.
3. The other restriction that can be loosened is that the random factor $\varepsilon_{1i}$ determining survey and item response is a single factor in the single-index model, hence perfectly correlated with itself. In the double-index model $\varepsilon_{1i}$ and $\varepsilon_{2i}$ can be allowed any degree of correlation by including in the item response probit an estimate of $\varepsilon_{1i}$ constructed from the results of the survey response probit. The expected value of $\varepsilon_{1i}$ given that $\beta_1 X_{1i} + \varepsilon_{1i} \geq 0$ can be constructed using the inverse Mills' ratio. However, when this was done, the coefficient estimates were not significant at standard levels.
4. The critical values $c_1$ and $c_2$ are normalized to zero in these procedures.
5. For the survey respondents, $\lambda_1 = E(\varepsilon_{1i}/\beta_1 X_{1i}+\varepsilon_{1i}\geq 0) = f(-\beta_1 X_{1i}))/(1-F(-\beta_1 X_{1i}))$, where f(.) and F(.) are the standard normal density and distribution functions, respectively. For income respondents, $\lambda_2 = E(\varepsilon_{2i}/\beta_2 X_{2i}+\varepsilon_{2i}\geq 0) = f(-\beta_2 X_{2i})/ (1-F(-\beta_2 X_{2i}))$.
6. The case of multiple independent selection criteria has been analyzed by Catsiapis and Robinson (1982).
7. The response index for the hours item indicates how many of five other unrelated items were answered, while the response index for income is based on two unrelated items.
8. There is evidently a high correlation between $\lambda_1$ and $\lambda_2$ in the income case. Although neither is significant in the regression shown in Table 4, an F-test for their joint significance results in $p = .06$. When either $\lambda_1$ or $\lambda_2$ but not both are included in an income regression, they are positive and significant with $p=.02$. This is somewhat puzzling, since $\lambda_1$ was not significant in the income item response probit.

Table 1

Ordered Probit Estimation Results

| Variables | Coefficient Estimates Income | Hours |
|---|---|---|
| Constant | 1.48** | 1.36** |
| Internal Medicine | -.24** | -.25** |
| Surgery | -.11+ | -.09 |
| Pediatrics | .25** | .28** |
| Obsetetrics/Gynecology | .02 | .02 |
| Radiology | .12 | .16 |
| Psychiatry | .26** | .17+ |
| Anesthesiology | .21* | .26* |
| Pathology | .19 | .20 |
| Other | .11 | .11 |
| Metropolitan (<1 million) | -.05 | -.02 |
| Nonmetropolitan | .20** | .24** |
| Female | -.05 | -.01 |
| Board Certification | .01 | .02 |
| AMA Membership | .12** | .15* |
| Foreign Medical Graduate | .12* | .17** |
| Hospital-based | .15* | .18* |
| Middle Atlantic | .04 | .08 |
| East North Central | .09 | .10 |
| West North Central | .10 | .10 |
| South Atlantic | .12 | .11 |
| East South Central | .05 | .02 |
| West South Central | .18* | .23* |
| Mountain | .23* | .17 |
| Pacific | .17* | .18* |
| Total calls | -.043** | -.043** |
| Years since M.D. | .019** | .019** |
| Ln (Years since M.D.) | -.546** | -.526** |
| $c_2$ | .388** | .066** |
| Log-likelihood | -4177.6 | -3248.4 |
| $\chi^2$(27 d.f.) | 380.0** | 347.9** |

Note: The reference categories are general family practitioners in New England metropolitan (>1 million) areas. The coding for categorical variables is 0 = no, 1 = yes.
+p<.10, *p<.05, **p<.01

Table 2

Binomial Probit Results

| Variables | Coefficient Estimates Survey | Income (1) | Income (2) | Hours (1) | Hours (2) |
|---|---|---|---|---|---|
| Constant | 1.40 | 1.77 | 0.75 | 2.36 | 1.40 |
| Internal Medicine | -0.26** | -0.10 | -0.09 | -0.09 | -0.08 |
| Surgery | -0.10 | -0.11 | -0.13 | 0.03 | 0.03 |
| Pediatrics | 0.29** | 0.01 | -0.01 | -0.12 | -0.19 |
| Obstetrics/Gynecology | 0.02 | -0.06 | -0.07 | -0.18 | -0.21 |
| Radiology | 0.23* | -0.31* | -0.23 | -0.63** | -0.65** |
| Psychiatry | 0.20* | 0.36* | 0.36* | -0.31 | -0.43+ |
| Anesthesiology | 0.33** | -0.16 | -0.08 | -0.40+ | -0.43+ |
| Pathology | 0.21 | 0.01 | 0.07 | -0.15 | -0.10 |
| Other | 0.09 | 0.08 | -0.08 | 0.03 | -0.03 |
| Middle Atlantic | 0.04 | -0.21 | -0.26 | -0.21 | -0.26 |
| East North Central | 0.07 | -0.12 | -0.14 | -0.24 | -0.28 |
| West North Central | 0.05 | 0.03 | 0.01 | -0.12 | -0.13 |
| South Atlantic | 0.08 | -0.03 | -0.04 | -0.21 | -0.26 |
| East South Central | -0.02 | 0.02 | 0.02 | -0.28 | -0.29 |
| West South Central | 0.22* | -0.27+ | -0.29+ | -0.40 | -0.43 |
| Mountain | 0.14 | 0.30 | 0.27 | -0.32 | -0.43 |
| Pacific | 0.15+ | -0.07 | -0.10 | -0.18 | -0.26 |
| Metropolitan (< 1 million) | -0.01 | -0.10 | -0.10 | -0.12 | -0.14 |
| Nonmetropolitan | 0.26** | -0.04 | -0.04 | -0.14 | -0.15 |
| Years since M.D. | 0.02** | 0.00 | 0.00 | -0.02 | -0.02 |
| Ln (Years since M.D.) | -0.54** | -0.25 | -0.27 | 0.09 | 0.12 |
| Female | 0.01 | -0.20* | -0.17+ | -0.11 | -0.05 |
| Board Certification | 0.02 | -0.01 | -0.03 | 0.09 | 0.11 |
| AMA Membership | 0.14** | -0.03 | -0.06 | 0.12 | 0.12 |
| Foreign Medical Graduate | 0.20** | -0.20* | -0.17* | -0.20 | -0.12 |
| Hospital-based | 0.19* | -0.03 | 0.10 | 0.01 | 0.07 |
| Total calls | -0.04** | -0.03** | -0.02** | -0.02* | -0.01 |
| Response index | - | - | 0.57** | - | 0.23** |
| Log-likelihood | -2792.4 | -1340.7 | -1311.6 | -417.7 | -395.8 |
| $\chi^2$(27 d.f.) | 358.65** | 110.22** | 168.48** | 43.84* | 87.8** |

Note: The reference categories are general family practitioners in New England metropolitan (>1 million) areas. The coding for categorical variables is 0 = no, 1 = yes.
+p<.10, *p<.05, **p<.01

Table 3

Income and Hours Regression Results

| Variable | (log)Net Income | (log)Total Hours |
|---|---|---|
| Constant | 8.144** | 3.832** |
| Internal Medicine | .172* | -.041+ |
| Surgery | .313** | -.074** |
| Pediatrics | -.061 | .033 |
| Obstetrics/Gynecology | .300** | .036 |
| Radiology | .287 | .102 |
| Psychiatry | .256+ | -.039 |
| Anesthesiology | .597** | .121* |
| Pathology | .225 | -.046 |
| Other | .338** | -.071** |
| Metropolitan (< 1 million) | .015 | .023 |
| Nonmetropolitan | -.066** | .055+ |
| Female | -.496** | -.089** |
| Board Certification | .293** | -.026+ |
| AMA Membership | .191** | .063** |
| Foreign Medical Graduate | -.158 | .082** |
| Hospital-based | .020 | -.037 |
| Middle Atlantic | .008 | .045 |
| East North Central | .157 | .024 |
| West North Central | .156 | .053 |
| South Atlantic | .122 | .060+ |
| East South Central | .295* | .108* |
| West South Central | .093 | .126* |
| Mountain | .155 | .037 |
| Pacific | .177 | .037 |
| Years since M.D. | -.070** | -.0070* |
| Ln (Years since M.D.) | 1.331** | .0584 |
| $\lambda_1$ | -.013 | .266** |
| $\lambda_2$ | .749 | -1.057* |
| $R^2$ | .2334 | .0816 |

Note: The reference categories are general family practitioners in New England metropolitan (>1 million) areas. The coding for categorical variables is 0 = no, 1 = yes.
+p<.10, *p<.05, **p<.01

Table 4

Actual and Predicted Sample Mean

| Item | Log (Net Income) | Log (Total Hours) |
|---|---|---|
| Respondents | | |
| - Actual mean | 11.3282 | 4.0016 |
| (std. error) | (.0224) | (.0063) |
| Full Initial Sample | | |
| - Predicted mean, regression with selectivity corrections | 11.3548 | 4.0032 |
| - Predicted mean, regression without selectivity corrections | 11.3660 | 4.0032 |