Information Collection Clearance

The Supporting Statement

Survey of Medical Examiners Who Certify the Physical Qualifications of Commercial Motor Vehicle Drivers

This is to request approval from the Office of Management and Budget (OMB) for a new survey on the role of medical examiners who certify the physical qualifications of commercial motor vehicle (CMV) drivers.

B. Collections of Information Employing Statistical Methods

1. Universe and Respondent Selection:

Background

Medical examiners who qualify CMV drivers are not conveniently available for use in sample surveys. They are included in unknown subsets of five different types of medical professionals based on their decision to perform this type of activity in their practice. They may be listed to various degrees in the rosters of a number of organizations but not necessarily identified as medical examiners, much less as to whether they examine CMV operators. Defined in this manner, medical examiners who perform this service constitute a hidden population relative to obtaining accurate information about their practice through a sample survey. Since there is no sample frame for hidden populations, attempts to study them using standard sampling and estimation techniques have often produced misleading results. When the NRCME program is fully implemented, the Registry itself will act as a sampling frame because it will contain the entire population of medical examiners who provide this service. Before that can be achieved, reasonable unbiased information is needed to validate the tasks performed by medical examiners which will permit the development of the National Registry. Given this, it is necessary to consider other approaches to obtaining useful information about medical examiners to allow this development to continue.

An approach which has gained much attention and produced promising results utilizes a collection of methods generally described under the term *network analysis*. The networks that are the focus of attention here are those social structures made up of family and friends, work partners, professional associates, acquaintances, and the organizations in which we participate. Anthropologists have studied the composition of these relationships, or social networks, in villages, towns and urban areas across the world. Social scientists have generally approached the work for two purposes. One is to identify the members of and the patterns of interactions among collections of individuals of various types such as kinship, work and friendship groups. The other purpose involves how to select respondents or participants for a research study by identifying individuals who know others or know about them. This latter activity is generally called *network sampling* or sometimes referred to as *snowball sampling* (Shensul et al. 1999).

More recently, however, the social network approach has been applied to more sophisticated problems; research on hidden populations is one of these areas. Here, the work has been directed at estimating the size of hidden populations and, of more interest to the present project, discovering the knowledge, beliefs and attitudes held by networks. In studying these areas, the effort has focused on how to estimate the parameters that describe social and professional networks and how these parameters behave under different sampling approaches and levels of missing data. The parameters of a network allow researchers to uncover information about members of the network even if they are not included in the sample. This reach into a network from a respondent is largely based on the concept of homophily. Homophily is the tendency for people to associate and develop friendships, marriages, work relationships, etc. with people similar to themselves. Various demographic characteristics such as age, gender, socio-economic class, ethnicity, education and occupation can provide boundaries around relationships. An individual's beliefs and values (e.g., religious, political) can also be influential. Individuals in homophilic relationships share common characteristics (e.g., beliefs, values, education) that make communication and relationship formation easier.

Parameters of most relevance to this study concern measures of centrality. Measures of centrality describe actors' (those asked about the network) positions in a network relative to others and in relation to the complete network. Centrality measures assess such network factors as degree, betweenness, closeness, radiality and integration. While these measures have technical definitions, their names are reasonably self explanatory except for degree. This measure relates to the number of individuals in the network known to an actor in the data collection process. This measure can be expressed as an individual's degree, the total degree of the network, or the network's average degree. It is further defined as being manifested in two ways, i.e., in-degree and out-degree. In-degree is measured as the number of other individuals who identify the actor and out-degree is the number of others who the actor identifies. To obtain a measure of in-degree, there is a need to collect in-depth information from all of the network members in the sample. Out-degree can be measured easily by asking respondents to report the number of individuals they know within the identified network boundary. Out-degree is the measure of greatest interest here because it is the type of information that can be obtained in the data collection processes to be used in this study. For example, one can ask a physician how many other physicians they know who are medical examiners who examine CMV operators to qualify them for an interstate commercial driver's license.

A number of researchers have studied the behavior of the degree measure in different circumstances. In this work, the main types of sampling strategies examined were network sampling and general survey sampling. Among this research, Salganik and Heckathorn (2004) showed that asymptotically unbiased estimates of degree can be obtained using a type of snowball sample. They found that estimated percentages of the population with specific traits were also asymptotically unbiased and that these estimates had this quality no matter how the initial members (seeds) of the network were selected.

Other researchers examined the effect of missing data. In his area, a study was performed that investigated stability of centrality measures when networks were sampled with various levels of missing data (Costenbader and Valente, 2003). The conditions investigated involved amounts of missing data ranging from 20% to 90%. Analyses of stability were performed by correlating the estimates taken from the full data set with those estimates based on the varying levels of missing data. The analyses were performed on 11 different sets of data and out-degree was found to have the most stability across the levels. The correlations were around .90 at 50% missing data and were still at .70 with 90% missing data. This is not surprising because the data provide an in-depth description of the network and are labor intensive to obtain. In-degree did not perform as well, but still had significant correlations (.60) with 50% missing data. This is encouraging because, as was mentioned, out-degree data can easily be obtained from respondents in a survey.

This indicates that under some circumstances researchers may still be able to use network data for which some data are missing to study network properties or create network-based interventions such as developing training and certification examinations. In other words, researchers who do not interview all members of a community or network may still be able to take advantage of some aspects of network theory and techniques to obtain useful information when the target group initially has the characteristics of a hidden population.

Sampling Strategy

In this study, we plan to obtain two types of samples. One will be obtained using a snowball type of approach while the other will be a random sample of medical examiners from organizations that are expected to have these individuals in their membership. Two samples are used so that more information will be available on this hidden population, and the results can be examined in different contexts. For example, we are confident that we can obtain unbiased results using the snowball sample but are not entirely sure of what an organizational survey may produce. If the results are similar, then there can be confidence that the estimation of the degree distribution is valid and will produce accurate estimates of the traits that are the focus of the task validation (see Estimation section). Should the results differ significantly, then there will be a need to investigate, explain and reconcile this discrepancy. One approach to doing this will involve asking the actors in each sample how many medical examiners they know who are members of the organizations involved and how many they know who are not. The degree distributions of the various subsets will be examined to determine if membership (or lack of membership) is a confounding factor. This approach is suggested through the concept of “critical multiplism” which is promoted by researchers and evaluators who argue that increased external validity of findings can be achieved by using multiple data sources even if some are flawed (General Accounting Office (GAO), 1992).

In 1992, the GAO issued a report that introduced a collection of approaches for generating coherent scientific evidence called Cross Design Synthesis. This report was critical of the flagship of medical scientific investigations, clinical trials, saying that they often had poor external validity. This was the case, the report said, because it was frequently difficult to apply findings to a broader patient population based on the methods of selecting the subjects into the trials. While internal validity of the approach was usually excellent, the weakness in relation to external validity was a serious flaw. The report argued for the broader use of non-experimental data to enhance the generalizability of findings. This meant, of course, there is a need to employ a range of methods to handle observational data so that findings can be based on a preponderance of evidence. The spirit of this approach focused on obtaining coherence from a complex program environment.

Underlying this view is the concept of “critical multiplism”. Here, its proponents point out that scientific results are often weakened because all aspects of the scientific approach are flawed and open to bias (Cook, 1985). It is argued that one can best approximate truth through a strategy of multiple approaches to demonstration of results in an attempt to control bias. In the course of doing this, one must adopt a critical focus relative to each approach. While this is not different from the concept of triangularization to achieve acceptable results, it is more of an elaboration. We feel it will be useful in this circumstance because medical examiners have not been systematically studied before and, as has been pointed out, they have some characteristics of a hidden population.

Estimation

The purpose of this survey is to determine the prevalence of tasks performed by medical examiners so as to validate earlier observational findings. As has been pointed out, the approach to this is to estimate the percentage of occurrence of a task as it is reported by the medical examiners. For example, the exclusion rule requires that a task 1) must be performed by a majority of examiners and 2) have an importance rating equal to or above a specific value to be used for testing for minimum competency. To provide data which allows a valid assessment of the tasks, it is necessary to obtain as unbiased estimates as possible of the percentages of task occurrence as reported by the medical examiners.

The best approach to this is to use the data describing network degree as an aid in making the estimates. Given the expected quality of the estimates of degree, the use of these with the data describing the occurrence of tasks will help adjust the prevalence estimates towards what actually occurs in the networks.

Guidance in conducting this type of analysis is provided by some recent research. Using survey data and measures of degree, researchers were able to develop models that combined the measures of interest to obtain estimates of the prevalence of traits in the network (Zheng et al., 2006). The researchers modeled network degree as count data in a non-homogenous Poisson framework. They used degree as a dependent variable in a Poisson regression with the occurrence of the traits as independent variables. The occurrence of the traits was estimated as the coefficients of the independent variables. We shall use the same approach with the occurrence of tasks as independent variables.

Sample Size and Response Rates. There are five critical subgroups in the population defined by a background as (1) an Advanced Practice Nurse, (2) a Doctor of Chiropractic, (3) a Doctor of Osteopathy, (4) a Medical Doctor, or (5) a Physician Assistant. We intend to identify 1,000 respondents for each of these five subgroups during an opt-in phase of the study. Surveys will be sent to 1,000 people from each subgroup who know about the proposed NRCME and the role of the survey in the project. We intend to follow up with sample members multiple times with a goal to achieve at least an 80% response in each subgroup. We feel that we will be able to realistically achieve this rate based on the saliency of the topic and the follow-up methods we will be using.

Only practitioners who have functioned as medical examiners will become constituents of the survey sample. Information released before the survey will inform constituents future medical examinations will only be performed by medical examiners listed on the NRCME . We expect that by informing sample members of the study purpose before they opt-in, the commitment to give responses will be higher than normal. Most importantly we intend to follow-up with weekly communications among those we know who have not responded.

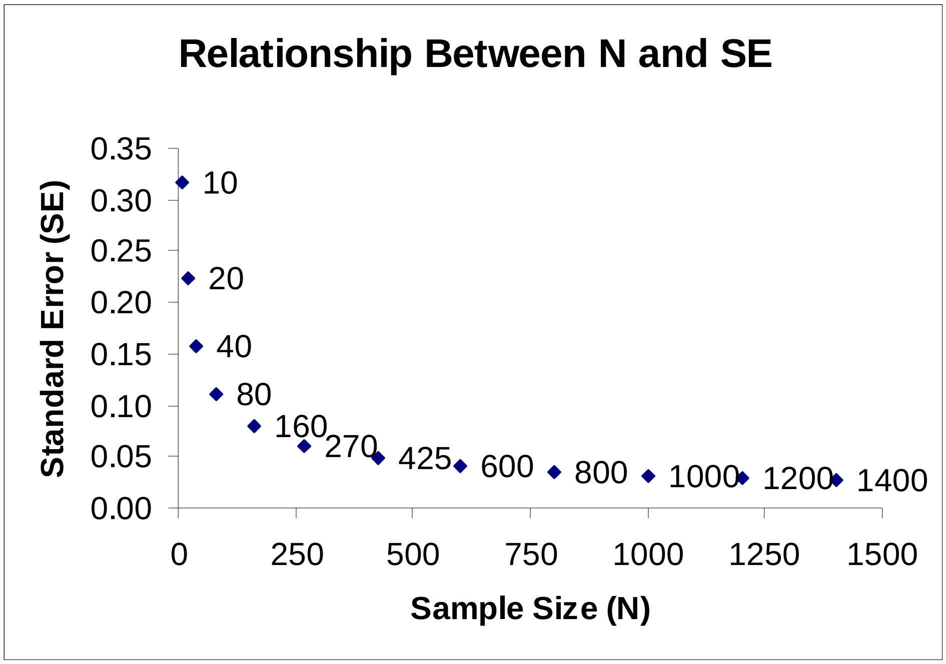

Figure 1

Erroneous deviations from population responses for a sample are inversely related to the square root of sample size, which Figure 1 displays. Error in observed survey responses would reach a practical minimum when approximately 400 responses are received. Our target response of 800 per subgroup should yield results in which survey response error is minimized for each subgroup.

We will calculate a standard error of the mean for each task so WIPT members can know when they risk making a false negative exclusion error. Task exclusion methodology will permit committee members to retain tasks within the margin of error for certification program content. Such a method will ensure only clearly peripheral tasks are excluded. Conversely, confidence will be high that surviving tasks over which certificants will be tested are truly critical.

2. <u>Procedures for Collecting Information:</u>

As indicated, the sample will be equally stratified among the five groups of professionals who are authorized to perform FMCSA CMV driver physical examinations. Within some of these smaller cohorts (e.g., Doctors of Chiropractic), the stratified sample may be the population of professionals. For larger cohorts (e.g., Medical Doctors) random sampling would be ideal. As discussed, one challenge in targeting survey respondents is that FMCSA does not currently know how many medical examiners are performing the FMCSA CMV driver physical examinations, the breakdown per medical profession or where these professionals are located.

Each cohort includes specialty practice areas where practitioners are more likely to perform the FMCSA CMV driver physical examination as part of their practice (e.g., within the Medical Doctor cohort, there are those who practice in occupational health settings). The proposed sampling will target a national distribution from these practice areas. However, the sampling will be completed from the general population of the professional group if there is an insufficient number of practitioners within the specialty areas. The goal is to increase the degree of accuracy and relevance of the responses. Only the initial survey will depend on targeting specialty areas. Subsequent surveys will be distributed to a sampling of the certified medical examiner population listed on the National Registry.

3. <u>Methods to Maximize Response</u>:

We will ensure that the survey sample of 5,000 includes those medical examiner sub-populations that perform the CMV physical, including 1,000 respondents from each of the five subgroups. Methods used to distribute the request to participate in the NRMCE survey include:

1. Word of mouth by medical examiners already contributing subject matter expert services to the NRCME program,
2. Professional association distribution through print and electronic publications,
3. Information prominently displayed on the NRCME Web site and other professional Web sites, and
4. Direct mailings to random samples of professionals in the high-probability specialty sub-populations to pre-screen that all participants perform the CMV driver physical examinations.

As of July 3, 2006, more than 2,400 survey participants have been identified.

The study will include an opt-in phase to ensure we start with 5,000 people who have completed physical examinations for CMV drivers and are willing to participate in the Medical Examiner Survey. We are more confident an opt-in sample will include persons who have at least some understanding of the study and are committed to following through by completing a survey. We will send postcards to mailing lists we expect include medical examiners. These individuals will be instructed to opt-in to receive a survey by registering their contact information at a Web site. These individuals will be added to the approximately 2,400 individuals who have already signaled their interest in receiving a survey until we have identified 1,000 willing recipients for each of the five subgroups.

Just prior to distribution of the survey, participants will be sent a postcard informing them when the survey will be sent, who to contact if they do not receive it, and encouraging them to complete the survey. Participants will also receive an electronic version of the message by e-mail.

Each survey will be sent in an envelope that is personally addressed to the medical examiner. This envelope will prominently display the NRCME banner and FMCSA logo to identify clearly the Federal source of its contents. The cover letter accompanying the survey will describe the proposed NRCME program and encourage each respondent to complete the survey. This cover letter will also be signed by a senior FMCSA Agency official. Survey introductory information will include a description of the typical time it should take to complete the survey, a support contact name, an email address and telephone number to answer recipients' questions as well as a confidentiality pledge. Pre-addressed and postage-paid envelopes will be included with the survey.

After the survey is distributed, the following intensive methods will be implemented to maximize response rate:

1. Weekly e-mail reminders encouraging prompt and complete responding until we have received at least 800 completed surveys from each of the five subgroups,
2. A mailed reminder; this letter will include a request for those not completing the full survey to respond to demographic questions,
3. Regular promotion of survey completion through the same professional and contact resources (e.g., listservs) that have been used to identify volunteers for the survey.

As indicated above, in an effort to capture information on non-respondents and assess non-response bias, the post-survey distribution reminder letter described above will include the same demographic questions as the survey. If participants do not intend to return the entire survey, they will be asked to respond to each demographic question contained in the letter. This way we hope to assess and account for nonresponse bias to the extent possible.

Intra-class correlation values will also help to provide an answer to the theoretic question, "Would other samples from the population have responded similarly to survey items?" Intra-class correlation values typically exceed .90 when the expected number of respondents reacts to a survey. Values this high are associated with high probability for similar responses from a survey sample and theoretic responses of other samples from the same population.

4. <u>Testing of Procedures:</u>

There was no pilot test of the survey instrument since the test development professionals involved in the Role Delineation Study have a well-documented, excellent record of past performance in conducting numerous role delineation studies.

5. <u>Contacts for Statistical Aspect of Data Collection:</u>

Kaye Kirby, Federal Motor Carrier Safety Administration, Program Manager, 202-366-3109

Glenna Tinney, Axiom Resource Management, Inc., Onsite Project Manager, 202-366-0549

Robert C. Shaw, Jr., PhD, Applied Measurement Professionals, Inc. (AMP), 913-495-4467

ATTACHMENTS

A. 49 U.S.C. 31149
B. SAFETEA-LU, Section 4116, Public Law 109-59, 119, Stat. 1726-28 (August 10, 2005)
C. 49 U.S.C. 31136
D. 29 CFR 1607.14
E. 49 CFR 390.5
F. Section 515 of Public Law 106-554
G. Cover Letter and Survey Instrument
H. FMCSA's 60-day notice requesting comments (Federal Register notice, dated September 29, 2005 (70 FR 56964))
I. Comment from the 60-day notice and FMCSA response thereto
J. Comment from the 60-day notice and FMCSA response thereto
K. FMCSA's 30-day notice requesting comments

References

Cook, T. D., (1985), "Post-Positive Critical Multiplism." In Shotland and Mark (Eds.), Social Science and Social Policy. Beverly Hills, CA, Sage.

Costenbader, E. and Valente, T. (2003), "The Stability of Centrality Measures When Networks are Sampled," Social Networks, 25, 283-307.

Raymond M.R. (2001). Job analysis and the specification of content for licensure and certification examinations. Applied Measurement in Education. 14(4), 369-415.

Salganik, M. J. and Heckathorn, D. D. (2004) "Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling." Sociological Methodology, 193-239.

Schensul, J., LeCompte, M., Trotter, R., Cromley, E. and Singer, M. (1999), Mapping Social Networks, Spatial Data and Hidden Populations, Beverly Hills, CA, Sage.

U.S. General Accounting Office, "Cross Design Synthesis: A New Strategy for Medical Effectiveness Research," March 1992, GAO/PEMD-92-18.

Zheng, T., Salganik, M. J. and Gelman, A. (2006) “How many people do you know in prison?: Using overdispersion in count data to estimate social structure in networks.” Journal of the American Statistical Association, to appear.