

# 36-303 Sampling, Surveys & Society

## Homework 02 Solutions

February 26, 2012

### **Part (a): Groves 2.2, p. 65**

Coverage error is due to the gap between the sample frame and the target population. Sampling error is due to the gap between the sample and the sample frame. Coverage error exists before any sampling is done, while sampling error results from the act of sampling, in that the sample is only a subset of the sample frame (i.e. not a census.)

### **Part (b): Groves 3.8, p. 96**

There is not one right answer. An acceptable answer:

**Undercoverage** would arise if not all of the clinic's patients had a record in the cabinet. If some records were older than two years old, and those patients had not visited since, they would constitute **ineligible units**. **Duplication** could occur if one person had multiple dossiers/background sheets instead of a single consolidated dossier, perhaps due to clerical oversight. **Clustering** could occur if, for example, two people named John Smith were lumped into the same dossier even though they had distinct identities. **Complicated mappings** are just compositions of the above.

### **Part (c): Groves 3.4, p. 95**

Again, there is not one right answer. An acceptable answer:

- duplication: some farms could be larger than a square mile and thus be sampled twice (by two different grid cells).
- ineligible units: some square mile cells could contain no farmland.
- clustering: there could be (pieces of) multiple farms within a single grid cell.

## Part (d): Groves 5.6, p. 179

- (a) Disagree; there is nothing inherent in the CAPI paradigm that concerns sample frame, so no generalization could be made about coverage error.
- (b) Agree: the sample frame of web-based surveys fails to capture households without internet access
- (c) Disagree: increasing the number of responses in Web-based surveys may help reduce sampling error but still fails to address coverage error. After some point, the incremental return on additional web-based responses will be small but the quality of the estimates may still be shoddy.
- (d) Disagree: it is difficult to assess nonresponse rates in surveys for which recruiting is done via banners, because it is impossible to know how many people see the banner and don't click on it.

## Part (e)

1. Coverage error: fence; facebook; email. Fence will fail to reach some CMU undergrads (undercoverage) while also reaching some people outside the target population (overcoverage). Facebook will similarly undercover but will not overcover. Email will do neither.

2. Nonresponse error: facebook; email; fence, or email; facebook; fence. It is clearly harder to ignore a person than an electronic solicitation. Which of facebook or email is harder to ignore is almost entirely subjective.

3. Measurement error: fence; facebook/email. If questions are displayed consistently and in writing, there is less chance for misinterpretation than in person; less chance for erroneous entering of answers by an interviewer; and less chance of dishonest response. Again, I don't see a clear distinction between facebook and email solicitations in this regard.