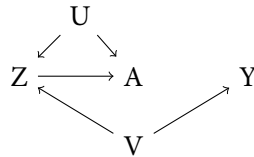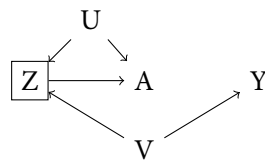# Homework 1

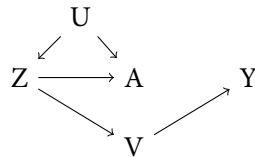Hard copy due January 31st, 2012 at the start of class

1   DAGs

For this question you will use the following causal directed acyclic graph:
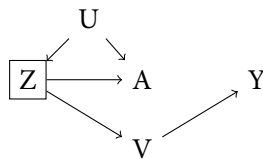

1. List all of the paths from A to Y and identify those paths as causal or noncausal.

2. If the above causal DAG is correct, would you expect there to be an association between A and Y?

3. Suppose that we control for Z (either in a regression or by running our analysis within levels of Z). In this analysis, would you expect there to be an association between A and Y?

4. Suppose now that we flip the direction of the arrow from V to Z, so that $Z \rightarrow V$. Would you then expect to see an association between A and Y in this revised DAG?

5. Suppose in this revised DAG we now control for *Z*. Would you expect to see an association between *A* and *Y*?

6. Add one arrow to the original DAG that doesn't point into or out of *A* but that does ensures that there will be an association between *A* and *Y* regardless of whether or not you control for *Z*. What arrow is that?

2   Experimental design

For this question, you should refer to "Monitoring Corruption" by Ben Olken.

1. In Table 2 of the Olken paper, he presents results from a Probit regression of the treatment on various village characteristics. What is this analysis intended to show? What untestable assumption is he attempting to build support for?

2. Note that Olken finds poverty and mountainous regions are correlated with the audit treatment, but that there are no such correlations for the invitation or comment treatments. What feature of the randomization might account for this?

3. Explain why Olken chooses to perform the randomization in the way that he does (in terms of the differences between audits and invitations). Which of the assumptions we have discussed are addressed by this randomization scheme?

3   Natural experiments

In "The Observer Effect in International Relations: Evidence from a Natural Experiment," Susan Hyde makes the argument that the placement of election observers in the 2003 Armenian election are "as-if" randomized. This problem will ask you to evaluate these claims and think about how violations of these claims might change the results of the paper.

1. Briefly (in a few sentences) summarize Hyde's results.

2. Suppose that Hyde is interested in the average treatment effect on the treated. Suppose further that her claims about randomization were false and election observers were more likely to monitor polling stations with higher support for the opposition. Using the concepts and notation from our class so far, make a detailed argument about what this would mean for her estimates. Would her *prima facie* effect correctly estimate, underestimate, or overestimate the ATT?

3. Now suppose election observers instead were more likely to monitor the incumbent strongholds. How would this affect her estimates?

4. What analysis in the paper would relieve these concerns about monitor placement being related to incumbent performance?