The rules of engagement for this final exam are the same as for the midterm. You may only use R. You may use canned random number generators, statistical distributions and simple functions such as *summarize*. You can take up to 24 hours to take the exam. Any references you consult will cost you 10 points. R-Help, class notes, this year's section notes, code and problem sets are free. Past years' code, exams and problem set answers are off limits completely. My number is 925-360-6473. But better try email as well. Also, I am asking you to study something that you do not know the correct answer to from class. I am trying to get you to use a Monte Carlo to help you figure out "whether you have a problem" or not. Just follow the instructions and tell me what you learn. I will give you the full writeup in the answer key with the right references.

This problem is designed to help you visualize the consequences of measurement error in your dependent and independent variables. One way to explore the consequences of measurement error on your estimated coefficients is to figure this out analytically, which is the preferred way (which you can look up in any textbook). Another approach is to conduct a Monte Carlo experiment, where you control the DGP. Here we are going to explore what happens to the least squares estimator for varying degrees of measurement error in the left hand and right hand side variables. Assume that you have the following population model:

$$y_i^* = \beta_o + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \varepsilon_i \tag{1}$$

Assume that $\varepsilon_i$ is distributed i.i.d. standard normal ($N(0,1)$), and $x_{1i}$ and $x_{2i}$ are both drawn from a uniform with support $[-200, 200]$. $\varepsilon$ is uncorrelated with the $x$ variables. Assume that $\beta_o = 1$, $\beta_1 = -0.75$, $\beta_2 = 0.75$. Set your seed to 22092008.

1. Generate a random sample of 100 observations for $x_{1i}$, $x_{2i}$ and $\varepsilon_i$.
2. Using the $x_{1i}$ and $x_{2i}$ from the step above and a 100 element vector of disturbances, generate the $y_i$. Estimate the three $\beta$ coefficients using least squares. Calculate the difference between the true $\beta$ and the estimated coefficient for each of the three $\beta$ coefficients.
3. Now assume that your $y_i$ is measured with error. It is in fact $y_i + r_i$, where $r_i$ is drawn from a random normal distribution with mean zero and variance $\sigma^2$. Using the $y_i$ from the previous step, add the measurement error $r_i$ to them and use this measured with error dependent variable as your outcome. Estimate the three $\beta$ coefficients using least squares. Do this for $\sigma^2 = [1\ 10\ 100]$.
4. Now assume that your $x_{2i}$ is measured with error. You observe $x_2^* = x_{2i} + r_i$, where $r_i$ is drawn from a random normal distribution with mean zero and variance $\sigma^2$ (Just reuse the $r_i$ from the previous step.). Use the $y_i$ from the first step (the one measured without error) and replace $x_{2i}$ with $x_2^*$ in your estimation. Estimate the three $\beta$ coefficients using least squares on your data. Do this for $\sigma^2 = [1\ 10\ 100]$.
5. Repeat step 4 exactly, only now assume that your measurement error is not symmetric, but always positive. Simply take the absolute value of $r_i$ (agin using the $r_i$ from above) before generating your $x_2^*$. Only do this for $\sigma^2 = 100$.
6. Repeat step 4 exactly, only now assume that your measurement error is not symmetric, but always negative. Simply take the absolute value of $r_i$ and multiply it times (-1) before generating your $x_2^*$ (again using the $r_i$ from above). Only do this for $\sigma^2 = 100$.

Repeat the above steps 10,000 times (set the seed only the first time, not each time) and calculate the bias for $\beta_o$, $\beta_1$ and $\beta_2$ for each setting and fill in the table below. What have you learned from this exercise that you did not know before? Was there anything surprising?

Table 1: Final Exam Results

| | Bias $\beta_o$ | Bias $\beta_1$ | Bias $\beta_2$ |
|---|---|---|---|
| Step 2 | | | |
| Step 3 ($\sigma^2 = 1$) | | | |
| Step 3 ($\sigma^2 = 10$) | | | |
| Step 3 ($\sigma^2 = 100$) | | | |
| Step 4 ($\sigma^2 = 1$) | | | |
| Step 4 ($\sigma^2 = 10$) | | | |
| Step 4 ($\sigma^2 = 100$) | | | |
| Step 5 | | | |
| Step 6 | | | |

Thank you all for paying attention all semester and not throwing hard objects at me for my bad jokes. I hope you get as much pleasure out of applying these tools as I do throughout your career.