**Temple University**

**Department of Economics**

**Economics 8009**

**Problem 1.** Consider the classical linear regression model except that one of the regressors is of the form This regressor vanishes in that as withprobability one. Show that the matrix is singular.

**Problem 2.** Define Px = X(X’X)-1X’ and Mx = I – Px. Let be the OLS estimator and be the least squares residuals for the model .

1. Shoe that the least squares estimator is an orthogonal projection of y onto the column space spanned by X.
2. Given your answer to (a), state Pythagorus’ theorem in terms of y, X, , and .
3. Use Px and Mx to state Pythagorus’ theorem in terms of only y.

**Problem 3.**  The data file available on the syllabus contains real data on participation in 401K plans in different organizations. The variables we will be interested in are PRATE, the *participation rate*, the percentage of employees enrolled in the company 401K plan, MRATE is the *matching rate*, the rate at which the employee contribution is matched by the employer, and AGE, the age of the employer’s 401K plan.

For example, Ordinary State University faculty typically contribute 10% of their salary to their retirement plan, which OSU matches with an amount equal to 10.5% of their salary, which makes for a matching rate of 1.05 or 105%. However, employees can often opt out of the retirement plan, and employers can contribute at different matching rate levels, implying that a wide range of outcomes is possible for the matching rate variable.

For this question, we will ask whether or not the level of participation in a 401K plan is affected by the matching rate.

a. Download the file from the web and into your software of choice. Generate a scatter plot

with PRATE on the *Y* axis and MRATE on the *X* axis. Do the data seem to suggest that the

participation rate will increase or decrease with an increase in matching rate? Motivate

your answer.

b. Do you see anything unusual about the data? Explain what might be the cause and how it

might affect your results.

c. Run a regression of PRATE on MRATE; report your results.

d. Generate a plot of the fitted regression line with a scatter plot of the data.

e. What is the interpretation of the intercept in this regression?

f. Now estimate the coefficients in the regression equation that includes both MRATE and AGE on the RHS. Report your results. How do you interpret the results?

g. Comparing the two regressions, with and without AGE, are the estimates of the MRATE coefficient different? Explain why or why not.

h. What does the AGE coefficient suggest about the behavior of employees or their beliefs about employer 401K plans?

**Problem 4** You are interested in the linear model

.

With the help of your research assistant, you have assembled the following data:

1. How many observations do you have?
2. What is the inverse of X’X?
3. What is the sum of the observations on y?
4. What are the least squares estimates of the model coefficients?
5. What is the residual sum of squares?
6. What is the standard error of your estimate for ?