

Exam #1

Economics 435

Spring 2003

1 Approval ratings

As you may know, the Premier of British Columbia, Gordon Campbell, was arrested in Maui last month for drunk driving. A surprising piece of news was reported by CBC last week: according to survey data, Mr. Campbell's approval rating *increased* by 5 percentage points following his arrest.

For some technical background, the relevant pair of surveys was conducted by McIntyre & Mustel, in December 2002 and February 2003. Mr. Campbell's arrest was around the beginning of January 2003. Each survey was conducted by contacting a random sample of 500 British Columbians and asking them the following question:

Do you generally approve or disapprove of the performance of Gordon Campbell as premier of the province?

The "approval rating" is simply the fraction of respondents who approve of Mr. Campbell's performance. In December 2002, exactly 35% of respondents approved of Mr. Campbell's performance, while exactly 40% of respondents in February 2003 approved of his performance.

Let p_d be the fraction of BC residents (not just those in the sample) who approved of Mr. Campbell's performance in December, and let p_f be the fraction in February who approved. Let \hat{p}_d be the average approval rating in the December sample, and let \hat{p}_f be the average approval rating in the February sample (i.e., $\hat{p}_d = 0.35$, $\hat{p}_f = 0.40$).

- a) Construct an estimate of the standard deviation of \hat{p}_f .
- b) Use this estimate to construct an asymptotic 95% confidence interval for Mr. Campbell's February approval rating. If you are not confident in your answer to the previous question, be sure that I can see what you are doing in this question if you want partial credit.
- c) We would like to get an idea how strong the evidence is for an actual increase in Mr. Campbell's level of approval, and how much is simply because of the relatively small sample. Conduct an asymptotic hypothesis test, at the 5% level of significance, of the null hypothesis that $p_f = p_d$ against the alternative that $p_f \neq p_d$.
- d) Suppose that you were to combine the two samples, and run an OLS regression of whether a person approved or disapproved on whether the person was asked in February. Define the two binary/dummy variables you would use for this regression, and explain how you would interpret the coefficients.
- e) Find the OLS estimates for this regression.

2 Estimating demand

You are an econometrician working for BC Hydro, and your boss wants your help on a project to estimate the elasticity of demand for electricity. In order to avoid the problem of endogeneity, BC Hydro randomly varies prices across different times and locations. Your boss estimates the following regression model:

$$\ln Q = \beta_0 + \beta_1 \ln P + u \quad (1)$$

- a) Assuming that your model is correct, and that β_1 is negative, draw this demand curve on a graph. Be sure to display the demand curve's properties. In particular, what is the quantity of electricity demanded when electricity is free?
- b) Suppose that $\hat{\beta}_0 = 7.689$ and that $\hat{\beta}_1 = -0.2$. By how much does quantity demanded increase when the price goes down by 1%?
- c) You believe that the elasticity of demand is sensitive to the weather. In particular, you think people are willing to pay a much higher price when it is hot enough. Describe an empirical strategy to address this issue. In particular, define the additional variables you would use (if available), define the regression you would estimate, and explain what hypothesis tests you might perform in order to determine whether demand is weather-sensitive.

3 Outliers

Two researchers are interested in estimating the relationship between height and intelligence, because they have hypothesized that taller people get less oxygen to their brains, and thus are not as smart as shorter people. Both researchers believe that the relationship is linear and takes the form of:

$$IQ = \beta_0 + \beta_1 height + u \quad (2)$$

where assumptions SLR1-SLR4 hold. The researchers are concerned about the impact of outliers on their results.

- a) Coauthor A believes that we should drop all observations with heights above 7 feet or below 5 feet. What is the cost to doing so? Will the OLS estimators be biased or unbiased? If biased, what is the likely direction of the bias? Explain your answer.
- b) Coauthor B believes we should drop all observations with IQ's above 150 or below 90. What is the cost to doing so. Will the OLS estimators be biased or unbiased? If biased, what is the likely direction of the bias? Explain your answer.

4 The augmented Solow growth model

Recall from your problem set that the simple Solow growth model performs poorly along several dimensions. Mankiw, Romer and Weill (MRW) develop what they call an "augmented" Solow growth model that incorporates human capital. Their theoretical model implies that

$$\ln\left(\frac{Y}{L}\right) = \text{Intercept} + \frac{\alpha}{1 - \alpha - \beta} \ln s_k + \frac{\beta}{1 - \alpha - \beta} \ln s_h - \frac{\alpha + \beta}{1 - \alpha - \beta} \ln(g + n + \delta)$$

where s_k is the rate of investment in physical capital, s_h is the rate of investment in human capital, g is the growth rate of (total factor) productivity, n is the population growth rate, and δ is the depreciation rate. In the model, the parameters α and β will also correspond to capital's share of output and human capital's share of output, respectively. It is assumed that $\alpha > 0$, $\beta > 0$, and $\alpha + \beta < 1$.

As before they have collected the variables $LOGYL$, $LOGS$, $LOGH$, and $LOGN$ (corresponding to the model variables $\ln \frac{Y}{L}$, $\ln s_k$, $\ln s_h$, and $\ln(g+n+\delta)$), and estimate the reduced-form regression:

$$LOGYL = \lambda_0 + \lambda_1 LOGS + \lambda_2 LOGH + \lambda_3 LOGN + u$$

Let $\hat{\lambda}_0, \dots$ be the OLS estimates of these coefficients.

- a) Identify the four testable implications (one each on the sign of λ_1 , λ_2 , and λ_3 , plus an additional restriction on a linear combination of these three coefficients) of the augmented Solow model for the OLS coefficients. State each implication as a null hypothesis/alternative hypothesis pair.
- b) Solve for α and β as a function of the λ 's. Use this to construct a formula for $\hat{\alpha}$ and $\hat{\beta}$, a pair of consistent estimators of α and β in terms of the $\hat{\lambda}$'s.
- c) You estimate the OLS regression and get the following results:

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 82 DF	P-VALUE
LOGS	0.51739	0.2600	1.990	0.050
LOGN	-0.61219	0.5362	-1.142	0.257
LOGH	0.92054	0.9586E-01	9.603	0.000
CONSTANT	3.7442	1.402	2.670	0.009

Calculate $\hat{\alpha}$ and $\hat{\beta}$. Is $\hat{\alpha}$ close to the theoretically-suggested value of 1/3?