Analysis of Economics Data
Chapter 8: Bivariate Case Studies

© A. Colin Cameron
Univ. of Calif. Davis

November 2022

CHAPTER 8: Bivariate Case Studies

1. Health Outcomes across Countries
2. Health Expenditures across Countries
3. Capital Asset Pricing Model
4. Output and Unemployment

Datasets: HEALTH2009, CAPM, GDPUNEMPLOY

8.1 Case Study 1: Health Outcomes across Countries

- Dataset HEALTH2009 has 2009 data for the 34 wealthy and relatively wealthy nations in the Organization of Economic and Community Development (OECD).
    - Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States.
- There is wide variation in **annual health expenditures per capita** and also in infant mortality.

Variable	Definition	Mean	St. Dev.	Min	Max
Hlthpc	Annual health expenditure p.c. (in US $)	3256	1494	923	7990
Lifeexp	Male life expectancy at birth (in years)	76.7	2.94	69.8	79.9
InfMort	Infant mortality per 1,000 live births	4.44	2.72	1.8	14.7

Life Expectancy and Health Spending per capita

- OLS regression yields

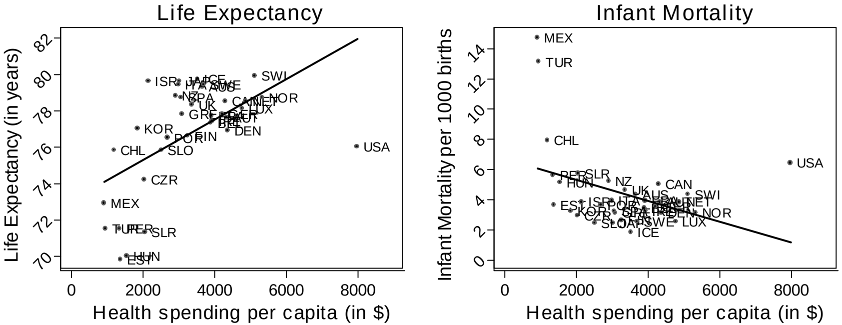
$$\widehat{Lifeexp} = \underset{(71.36)}{73.08} + \underset{(3.88)}{0.00111} \times Hlthpc, \quad R^2 = 0.320, \ s_e = 2.46, \ n = 34,$$

  where t-statistics based on default standard errors are given in parentheses.

- The relationship is **economically significant**
  - A \$1,000 increase in per capita health spending, a two-thirds of a standard deviation change, is associated with an increase in life expectancy of 1.11 years.
- The relationship is highly **statistically significant**, as $t = 3.88$.
- Here prior belief is that $\beta_2 > 0$ so do a one-sided test of $H_0 : \beta_2 \leq 0$ against $H_a : \beta_2 > 0$
  - $c = t_{32,.05} = 1.69$ so reject H_0 at significance level 0.05 since $t = 3.88 > c$.
  - or $p = \Pr[T_{32} > 3.88] = 0.000$ to three decimal places.

Infant Mortality

- The second panel additionally studies infant mortality.
- The U.S. has much worse outcomes than predicted by the model.



Further Details

- For these cross-section data with independence across observations it is standard to use **heteroskedastic-robust** standard errors.
- For life expectancy slope we obtain heteroskedastic-robust standard error 0.0004637 compared to default 0.000287
  - the t statistic falls to 2.40 from 3.88; still statistically significant at 5%.
- The plotted relationships appear to be **nonlinear** rather than linear
  - log-linear and log-log models are presented in Chapter 9.
- Bottom line: Country health outcomes improve on average with higher health spending
  - U.S. performs substantially worse than predicted.

8.2 Case Study 2: Health Expenditures across Countries

- Again use dataset HEALTH2009.
- Health expenditure is measured per capita, and income is measured using GDP per capita.
- There is considerable variation in GDP per capita, measured in current US dollars at current exchange rates, ranging from $13,807 for Mexico to $82,901 for Luxembourg, a small European country with population of half a million.

Variable	Definition	Mean	St. Dev.	Min	Max
Gdppc	GDP per capita (in US $)	33054	12918	13807	8290
Hlthpc	Health expenditure per capita (in US $)	3256	1494	923	7990

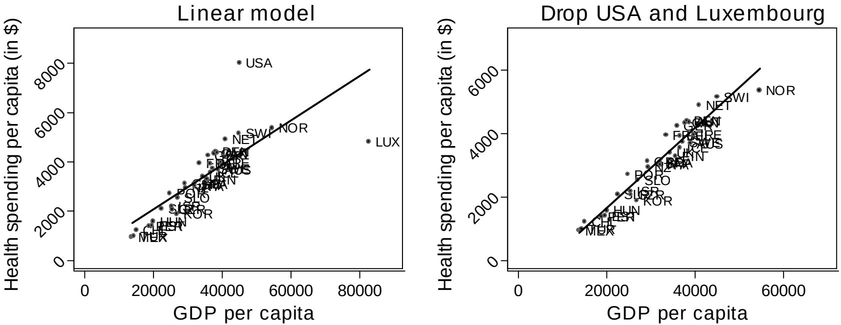
Health Spending per capita and GDP per capita

- OLS regression yields

$$Hlthpc = \underset{(0.63)}{285} + \underset{(6.99)}{0.0899} \times Gdppc, \quad R^2 = 0.604, \;\; s_e = 954, \;\; n = 34,$$

  where t-statistics based on default standard errors are given in parentheses.
- Slope coefficient estimate implies that an extra $1,000 in GDP per capita is associated with an $89.90 increase in per capita health expenditures.
- Relationship is highly statistically significant, as $t = 6.99$.
- Here prior belief is that $\beta_2 > 0$ so perform a one-sided test of $H_0 : \beta_2 \leq 0$ against $H_a : \beta_2 > 0$
  - again reject H_0.

- The U.S. has unusually high health expenditures
  - much higher than other countries and roughly \$4,000 more than predicted by the line.
- Similarly Luxembourg seems to be an outlier.
- The second panel drops the U.S. and Luxembourg
  - the slope coefficient increases from 0.0809 to 0.01267 and $R^2 - 0.928$..



Heteroskedastic-robust Standard Errors

- For these cross-section data with independence across observations it is standard to use heteroskedastic-robust standard errors
- For life expectancy slope we obtain **heteroskedastic-robust** standard error 0.0293 compared to default 0.0129
  - then t statistic falls to 3.08 from 6.99, but is still statistically significant at 5%.
  - the large change is due to large residuals for U.S. and Luxembourg.

8.3 Case Study 3: Capital Asset Pricing Model

- Dataset CAPM has monthly data from May 1983 to October 2013 on
  - returns to holding stock in Coca-Cola, Target and Walmart
  - one-month U.S. Treasury bill rate
  - market return = value-weighted return on all stocks listed on the NYSE, AMEX and NASDAQ.

Variable	Definition	Mean	St. Dev.	Min	Max
RM	Market return	.0091	.0456	-.2254	.1285
RF	One-month U.S. T Bill rate	.0035	.0022	.0000	.0100
RKO	Return on Coca-Cola	.0137	.0618	-.1909	.2266
RTGT	Return on Target	.0138	.0842	-.4781	.2673
RWMT	Return on Walmart	.0156	.0703	-.2698	.2644
RM-RF	Excess Market Return	.0055	.0456	-.2314	.1243
RKO-RF	Excess Return on Coca-Cola	.0102	.0616	-.1952	.2188
RTGT-RF	Excess Return on Target	.0103	.0842	-.4841	.2629
RWMT-RF	Excess Return on Walmart	.0121	.0702	-.2758	.2612

Theory of Capital Asset Pricing (CAPM)

- RF is the **risk-free interest rate (**one-month U.S. Treasury bill).
- RM is the **overall market return** on stocks.
- $(RM - RF)$ is the **market excess return** or the **equity market premium**.
- RA is the return on the **investment asset** A, here Coca-Cola.
- **CAPM** links the (excess) returns on individual investments to the market excess return

$$\mathrm{E}[RA_t - RF_t] = \beta_A \mathrm{E}[RM_t - RF_t].$$

  - β_A is the "**beta**" and is on average one across the market.
- Estimate by OLS

$$RA_t - RF_t = \alpha_A + \beta_A (RM_t - RF_t) + u_t.$$

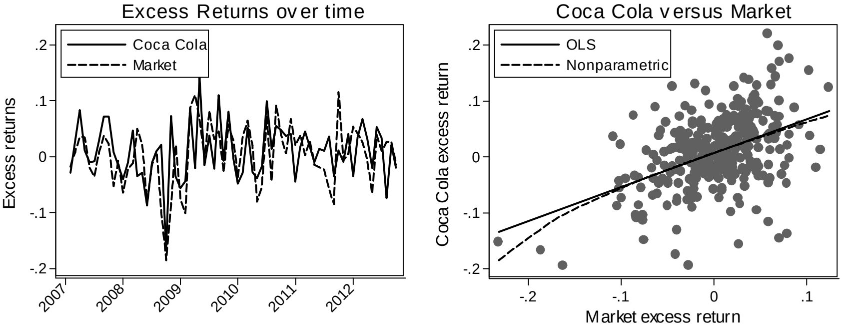
Estimated CAPM Model

- OLS regression gives fitted CAPM model for Coca-Cola (RKO)

$$(RKO - RF) = \underset{(0.00295)}{0.00681} + \underset{(0.0644)}{0.6063} \times (RM - RF), \quad R^2 = 0.201, \quad s_e = 0$$

  where default standard errors are given in parentheses.
- Slope coefficient is the stock's **beta**
  - statistically different from zero:
    $t = 0.6063/0.0644 = 9.41 > t_{0.025,352} = 1.967$.
  - statistically different from one: $t = (0.6063 - 1)/0.0644 = -6.11$.
  - value stock as beta lies between 0 and 1.
  - large companies such as Coca-Cola generally move less than the market as a whole, leading to $\beta < 1$.
- Intercept coefficient is the stock's **alpha**
  - a risk-adjusted measure of stock performance that measures the return in excess of that expected given the riskiness of the stock.
  - CAPM model in its purest form restricts $\alpha = 0$.
  - restriction rejected: $t = 0.00681/0.00295 = 2.31 > 1.967$.
  - furthermore the alpha is large in magnitude.

- For readability the first panel uses only the last 20% of the sample.
- The second panel uses all data from 1983 to 2013.



Robust Standard Errors

- Default standard errors assume error independence over time and homoskedasticity.
- For time series in general model errors may be correlated over time.
- For financial returns data, however, excess returns are intrinsically not forecastable if markets are efficient. So the error term should be uncorrelated.
  - here $\text{Corr}(e_t, e_{t-1}) = -0.039$ is close to zero.
- The **heteroskedastic-robust standard error** of the slope coefficient is 0.0770 (compared to default se of 0.0644).
- A **HAC standard error** that additionally controls for error correlation (see Chapter 12.1) is 0.0885.

8.4 Case Study 4: Output and Unemployment in the U.S.

- Dataset GDPUNEMPLOY has annual U.S. data from 1961 to 2019.
- *Growth* is the annual percentage growth in real GDP.
- *URATEchange* is the annual change in the percentage unemployment rate for the civilian population aged 16 years and older.
    - e.g. if unemployment rate increases from 5.3% to 6.5% then *URATEchange* equals 1.2.

Variable	**Definition**	**Mean**	**St. Dev.**	**Min**	**Max**
Growth	Annual % growth in real GDP	3.059	2.038	-2.537	7.237
URATEchange	Annual change in unemployment rate	-0.032	0.987	-2.143	3.530

Okun's Law

- Okun's law is that each percentage point increase in the unemployment rate is associated with an approximate two percentage point decrease in the GDP growth rate.
  - called **Okun's law** after Okun who first proposed it in a 1962 journal article
  - better term is "Okun's rule-of-thumb" as it is an empirical relationship rather than an ironclad law.

Estimated Model

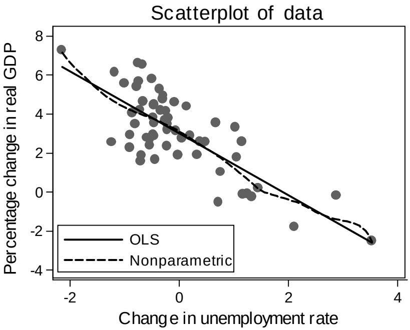
- OLS regression yields

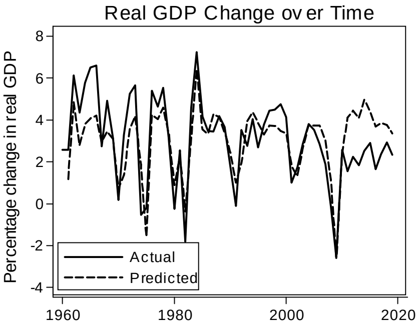
$$\widehat{Growth} = \underset{(0.162)}{3.008} - \underset{(0.175)}{1.589} \times URATEchange, \quad R^2 = 0.592, \quad s_e = 1.313,$$

  where default standard errors are given in parentheses.
- Slope coefficient is highly statistically significant with $t = -1.589/.175 = -9.09$.
- Test of Okun's law is test of $H_0 : \beta_2 = -2.0$ against $H_a : \beta_2 \neq -2.0$
  - $t = (1.589 - 2.0)/0.175 = -2.35$, so $p = \Pr[|T_{23}| \geq 2.35] = 0.022$.
  - null hypothesis is rejected at significance level 0.05.
  - so Okun's law is rejected by the data at 5%, though -1.59 is reasonably close to -2.0.

Prediction

- From second panel, output recovery from the 2008 global financial crisis is not as strong as predicted by the model.





Robust Standard Errors

- Default standard errors assume error independence over time and homoskedasticity.
- For time series such as this model error is in general correlated over time.
- A **HAC standard error** that additionally controls for error correlation (see Chapter 12.1) is 0.207 compared to the default standard error of 0.175.