

MODEL NUMBER 4 OF A “PUBLIC AWARENESS ANALYZED EXAMPLE”

JONATHAN A. PORITZ
MATH 156, FALL 2008
COLORADO STATE UNIVERSITY, PUEBLO

WHERE: This is found at

<http://www.soi.uzh.ch/research/wp/2008/wp0810.pdf>

WHAT: It's a graph from a paper comparing “income redistribution” in the United States and in Europe, written by two economists at the Socioeconomic Institute of the University of Zurich, Switzerland. Here it is:

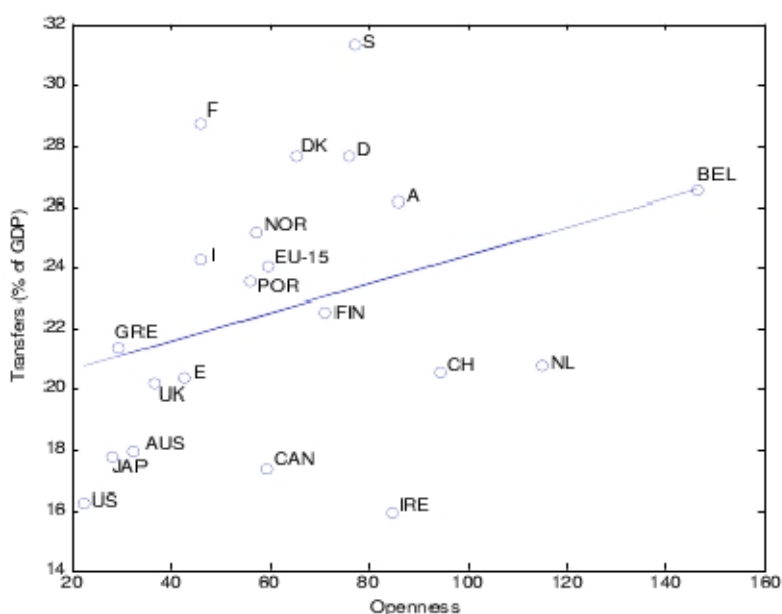


Figure 2a: Relationship $y = \hat{\alpha} + \hat{\beta}x$ between transfers y (% of GDP in 2003) and openness x , defined as sum of exports and imports 2007 over GDP in 2007; $\hat{\alpha}=19.609$, $\hat{\beta}=0.0473$, $R^2=0.1094$, t statistic is 1.433 (not significant). Data source: CIA World Factbook 2008.

ANALYSIS: Data was collected from various industrialized countries – the **observational units** were these countries – on two economic factors, the “transfers” and the “openness” – the **Variables** of (this part of) the study, both **quantitative** – expressed as fractions of the respective nation’s GDP. A **scatterplot** was produced, the coefficients of the **least squares regression line** were computed, the LSRL was plotted on top of the scatter plot, and the **correlation coefficient** (well, actually, it’s square, R^2) was noted in the caption of this figure. [The t -statistic, a measure of the reliability of the correlation coefficient in this dataset, was also noted in the caption; this is a consideration we will not cover in this course.]

CRITICISM: The scatterplot was reasonably clear, with good labels and title, and well-chosen window on the data. It is fairly clear upon looking at this plot that there is little useful information in the regression line, and the size of the correlation coefficient confirms this. Unless this paper is making some point which relies upon the *lack of association* between the two variables under study, there is little in this picture from which a conclusion can be drawn – in particular, if it is important for the argument of the paper that the association is **positive**, that would seem to be an almost entirely accidental feature of the data, given the small value of the correlation coefficient.