Draft1

Sam

December 11, 2014

Quantitative Analysis of the Impact of Education on Entrepreneurship

Data Sources

To complete a quantitative analysis of the impact of educational factors – specifically the encouragement of creatitivity, self-sufficiency, and personal initiative – on the country-level rate of entrepreneurship, we have drawn on publicly available data from the Global Entrepreneurship Monitor (GEM) National Expert Survey (NES) and Adult Population Survey (APS).

The Global Entrepreneurship Monitor project is an annual assessment of the entrepreneurial activity, aspirations, and attitudes of individuals across a wide range of countries. The countries covered have grown from a low of ten to covering 75% of world population and 89% of world GDP by 2013. The 2001-2010 multi-year APS data base, for example, contains over 1.3 million observations of a total of 85 countries. This data is often cited in academic research, including reports by the Organisation for Economic Co-operation and Development.

The NES is structured around nine entrepreneurial framework contitions evaluated by thirty-six experts in each surveyed country. These nine entrepreneurial framework conditions are included below (a minimum of four experts must be interviewed on each of the framework conditions).

1. Finance

2. Government policies

3. Government programs

4. Entrepreneurial Education and Training

5. Research and Development Transfer (R & D)

6. Commercial and professional infrastructure

7. Internal Market openness

8. Physical infrastructure and services

9. Cultural and social norms

The APS is a questionnaire distrubuted to a minimum of 2,000 adults in each participating country. The survey is created by a central GEM team, but the implementation is managed by national teams that compete for the role of implementing the survey. Experts are chosen based on experience and specialization and are expected to represent the entire country (including urban and rural areas).

Data Selection

After reviewing the available data, we decided to select the following variables for our analysis.

Source	Indicator
APS	Rate of Entrepreneurship (Intention, Nascent, Established)

Source	Indicator
NES	Way of Teaching
APS	Perceived Capabilities
APS	Fear of Failure / Perceived Opportunities
APS	Perceived Opportunities

The survey questions corresponding to these variables are as follows:

1. **Entrepreneurial Intention**: Percentage of 18-64 population (individuals involved in any stage of entrepreneurial activity excluded) who intend to start a business within three years.

2. **Nascent Entrepreneurship**: Percentage of 18-64 population who are currently a nascent entrepreneur, i.e., actively involved in setting up a business they will own or co-own; this business has not paid salaries, wages, or any other payments to the owners for more than three months.

3. **Established Entrepreneur**: Percentage of 18-64 population who are currently owner-manager of an established business, i.e., owning and managing a running business that has paid salaries, wages, or any other payments to the owners for more than 42 months.

4. **Way of Teaching**: In my country, teaching in primary and secondary education encourages creativity, self-sufficiency, and personal initiative.

5. **Perceived Capabilities**: Percentage of 18-64 population who believe to have the required skills and knowledge to start a business

6. **Fear of Failure**: Percentage of 18-64 population with positive perceived opportunities who indicate that fear of failure would prevent them from setting up a business.

7. **Perceived Opportunities**: Percentage of 18-64 population who see good opportunities to start a firm in the area where they live

The APS data represent the national-level average of values between 1 and 100 that were gathered as responses to surveys within each country. The NES data is calculated based on the average answer to a survey to the national experts. For each question, the experts surveyed had to state whether it is:

1. Completely false
2. Somewhat false
3. Neither true or false
4. Somewhat true
5. Completely true

(97. Do not know)
(98. Not applicable)

The process of importing, cleaning, and analyzing all of the data used in this project is fully documented and reproducable by referencing our GitHub repository.

Analytical Framework

To understand the impact of educational factors – specifically the encouragement of creativity, self-sufficiency, and personal initiative – on the country-level rate of entrepreneurship we built our model as outlined in the table below.

Variable	Indicator
Y Dependent Variable	Rate of Entrepreneurship (Intention, Nascent, Established)
X1 Explanatory Variable	Way of Teaching
X2 Control Variable	Perceived Capabilities
X3 Control Variable	Fear of Failure / Perceived Opportunities
X4 Control Variable	Perceived Opportunities

..........(Describe why we use the rate of entrepreneurship indicators we chose)

Data on our explanatory variable, Way of Teaching, is available on the GEM website for the years 2001-2010. To illustrate the distribution of this data, we have created the following histogram. As shown in the histogram, the largest density of scores is approximately 2.3, which lies between the "somewhat false" and "neither true or false" categories. This is not a whole number because the scores given by each of the 4 experts surveyed for each frameork condition are averaged to create a national-level number. As discussed later in this paper, this presents some difficulties in drawing substantive conclusions from this data. Namely, a large number of indifferent responses complicate the process of determining the influence we would like to analyze.
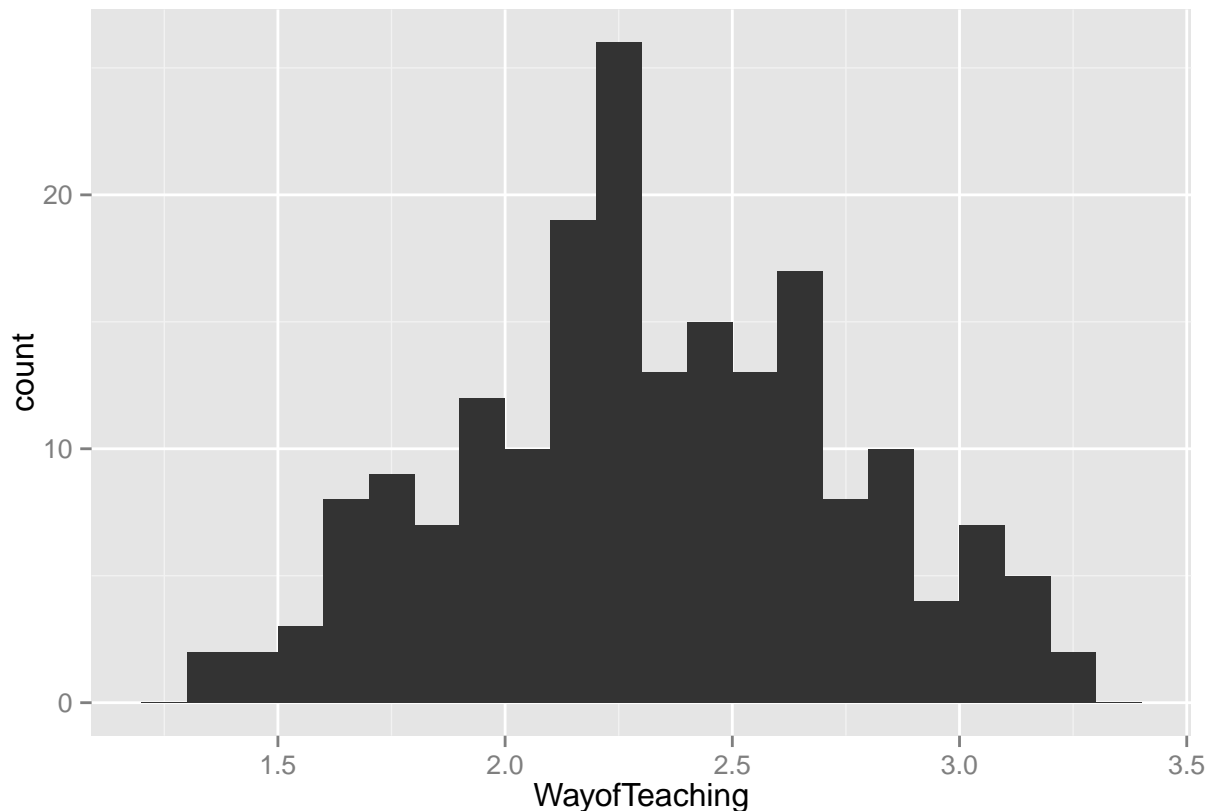
In addition to understanding the density of each recorded "Way of Teaching" indicator, it is also helpful to examine the differences in this indicator among OECD countries. To see this, we have created a figure showing the distribution of the average "Way of Teaching" value over the 2001-2010 period for the OECD countries included in the surveys. This figure is included below.

latex

Starting with this explanatory variable, we built three regression equations to examine the impact of our explanatory variable on the rate of entrepreneurship (as measured by survey responses on the percentage of nascent entrepreneurs, established business owners, and respondents with the intention of becoming an entrepreneur). These three regression equations are as follows:

1. $PercentEntreIntention = \beta_1 WayofTeaching + \beta_2 PerceivedCability + \beta_3 PerceivedOpportunities$
   $+\beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$

2. $PercentNascent = \beta_1 WayofTeaching + \beta_2 PerceivedCability + \beta_3 PerceivedOpportunities$
   $+\beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$

3. $PercentOwner = \beta_1 WayofTeaching + \beta_2 PerceivedCability + \beta_3 PerceivedOpportunities$
   $+\beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$

The interaction term $\beta_5 FearFailure * PerceivedOpportunities$ was added to the regression because the survey question regarding "Fear of Failure" was only proposed to survey respondents who indicated they do perceive opportunities to start a business.

Based on a review of academic literature and expectations from other data analysis, we expected to see a statistically significant positive relationship between self-confidence reinforcement in teaching and the dependent variables we selected as indicators of the entrepreneurship rate. Instead, we found that when controlling for country-specific effects, the Way of Teaching is not statistically significant and has a negative relationship with our explanatory variables. This relationship can be seen in the following regression output table.

Overall, however, if we omit country-specific effects, we can demonstrate that the Way of Teaching and the Percent Perceived Good Opportunities are statistically significant in the model indicated below.

$NationalLevelAveragePercentOwningaBusiness = -6.82 + 1.03 WayofTeaching + 0.10 PercentPerceivedCapability + 0.21 PercentFearFailure + 0.12 PercentPerceivedOpportunities - 0.005 PercentFearFailure * PercentPerceivedOpportunities$

We can show the relationship between the dependent and explanatory variables in the four simulations below. The R package Zelig allows us to set parameters for an explanatory variable and simulate 1000 trials of the above equation (more information on the Zelig package can be found on the program's CRAN page. We chose to simulate values in "Percent Perceived Capabilities" and "Way of Teaching" for this analysis. A regression output table has also been included below.

The simulations labeled "Unimputed" were created using only data collected from GEM. This simulation is not possible with missing data, so observations with missing data were omitted. In an attempt to increase the accuracy of this model, we used the R program Amelia to generate an imputed data frame. This program uses statistical inference to generate this data (more information on the Amelia package can be found on the program's CRAN page). Simulations titled "Imputed" were calculated using this imputed data frame.
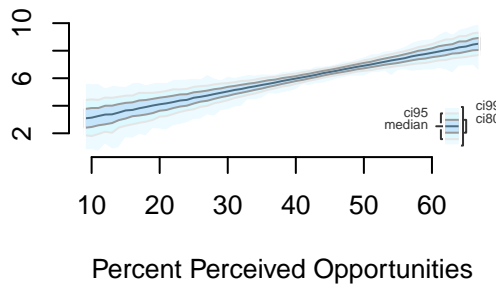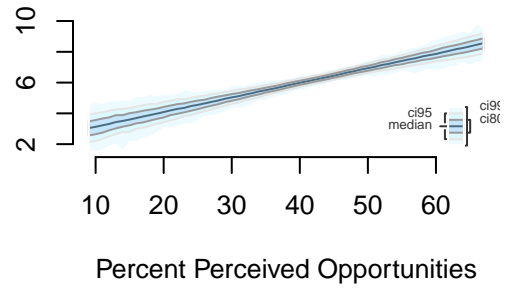
Unimputed (Equation 1)

Percent Entrepreneurship Owner

Percent Perceived Opportunities

Imputed (Equation 2)

Percent Entrepreneurship Owner

Percent Perceived Opportunities

Unimputed (Equation 1)

Percent Entrepreneurship Owner

Way of Teaching

Imputed (Equation 2)

Percent Entrepreneurship Owner

Way of Teaching
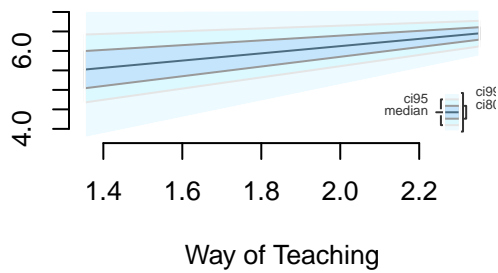
Table 3: Linear Regression Estimates

| | *Dependent variable:* | |
| | PercentOwner | |
	(1)	(2)
WayofTeaching	0.94*	1.03***
	(0.52)	(0.40)
PercentPerceived	0.09***	0.09***
	(0.02)	(0.02)
PercentFear	0.25***	0.21***
	(0.07)	(0.04)
PercentOpportunities	0.14**	0.12***
	(0.07)	(0.05)
PercentFear:PercentOpportunities	−0.005**	−0.004***
	(0.002)	(0.001)
Constant	−7.67***	−6.82***
	(2.78)	(2.00)
Observations	172	318
R^2	0.19	0.17
Adjusted R^2	0.17	0.15
Residual Std. Error	2.58 (df = 166)	2.71 (df = 312)
F Statistic	8.03*** (df = 5; 166)	12.38*** (df = 5; 312)

Note: *p<0.1; **p<0.05; ***p<0.01

Regression equation 1 is unimputed, regression equation 2 is imputed