

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 creativity, 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 conditions 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 distributed 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)
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)

Additional information on the GEM data used for this research can be seen in the descriptive statistics table included below.

% Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Fri, Dec 12, 2014 - 20:59:57

Table 2:

Statistic	N	Mean	St. Dev.	Min	Max
WayofTeaching	192	2.324	0.419	1.353	3.278
PercentFear	318	33.928	7.982	10	61
PercentNascent	318	4.329	2.505	0.400	15.400
PerceivedCapability	318	43.186	11.065	9	67
PercentOwner	317	6.414	2.931	0.400	28.000
PercentIntention	295	10.299	6.951	0.700	46.500
PercentOpportunities	318	33.931	15.168	3	71

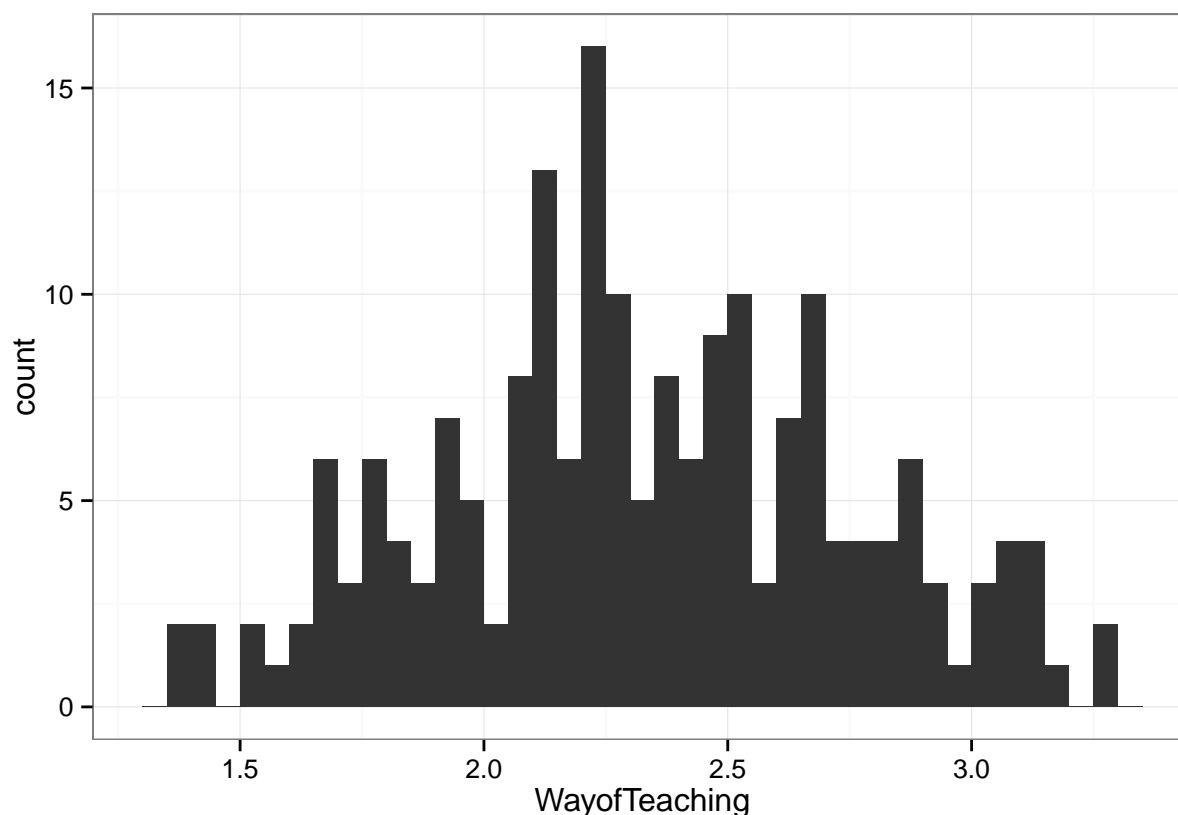
Analytical Framework

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.

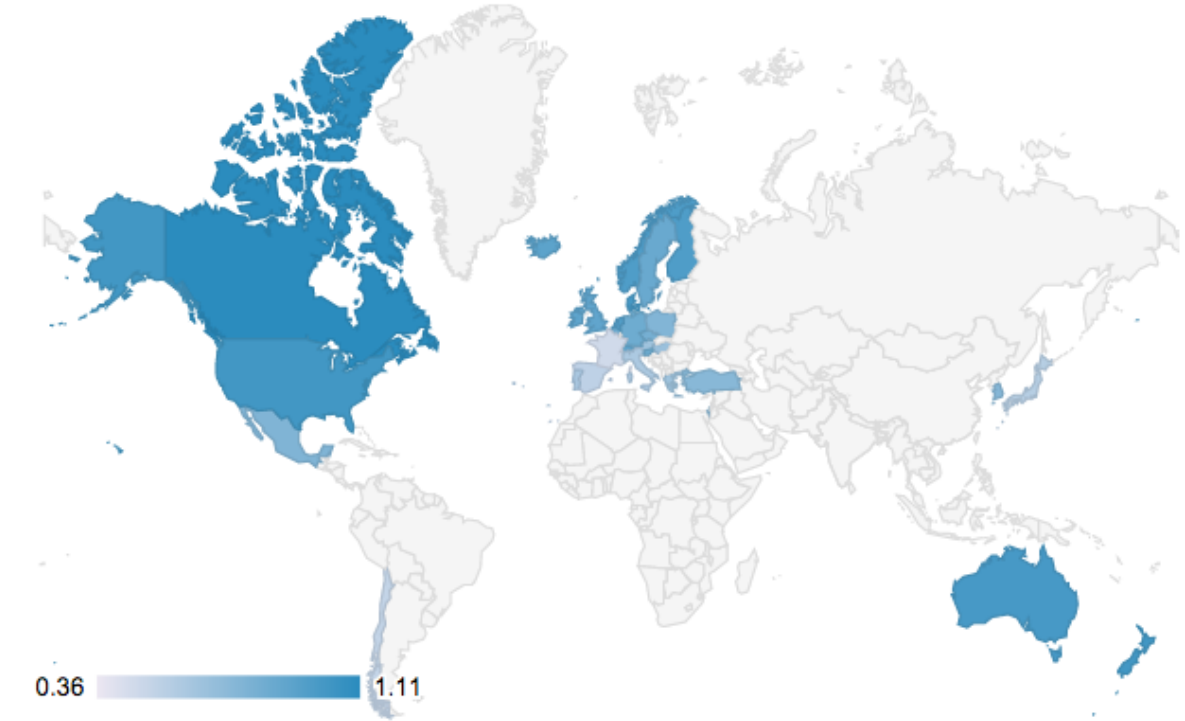


image:

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. $PercentEntreIntention = \beta_1 WayofTeaching + \beta_2 PerceivedCabability + \beta_3 PerceivedOpportunities + \beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$
2. $PercentNascent = \beta_1 WayofTeaching + \beta_2 PerceivedCabability + \beta_3 PerceivedOpportunities + \beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$
3. $PercentOwner = \beta_1 WayofTeaching + \beta_2 PerceivedCabability + \beta_3 PerceivedOpportunities + \beta_4 FearFailure + \beta_5 FearFailure * PerceivedOpportunities$

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 when using “Nascent Entrepreneur” or “Business Owner” as a dependent variable, has a negative coefficient in these regression equations. This relationship can be seen in the following regression output table. This regression output was generated using a simple least square regression model.

% Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Fri, Dec 12, 2014 - 20:59:59

To more clearly demonstrate the relationship between the “Way of Teaching” and our dependent variables, we have omitted the country-specific effects and generated simulations to show demonstrate the outcome of 1000 trials of our regression equation. Before doing so, however, we ran another set of regressions to determine which equation we would like to examine more closely. A regression output table with these results has been included below.

% Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Fri, Dec 12, 2014 - 20:59:59

Given the positive relationship between the “Way of Teaching” and “Entrepreneurship Owner” we decided to use this equation for our simulations. In an effort to address data availability issues, we also decided to run a second model including imputed data generated by the R program Amelia to generate an imputed data frame. This program uses statistical inference to generate imputed data (more information on the Amelia package can be found on the program’s [CRAN page](#)). Equation titled “Imputed” were calculated using this imputed data frame.

To generate the simulations, we decided to use the R package Zelig. This package allows us to set parameters for an explanatory variable and simulate 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 our explanatory variable “Way of Teaching” and “Perceived Capabilities” for this analysis. One important aspect of the Zelig package is that observations with omitted variables will be omitted from the calculation. For this reason, there is some variation between the coefficients for the Zelig model and the model in the preceeding regression output chart.

Given our data characteristics and the available models in the Zelig package, we used a “Least Squares Regression for Continuous Dependent Variables.” A regression output table and our simulation results are included below.

% Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Fri, Dec 12, 2014 - 21:00:01

Table 4: Linear Regression Estimates with Country-specific Effects

	<i>Dependent variable:</i>		
	Entrepreneurial Intention	Nascent Entrepreneur	Entrepreneurship Owner
	(1)	(2)	(3)
Way of Teaching	0.57 (0.98)	−0.04 (0.41)	−0.44 (0.67)
Perceived Capabilities	0.09* (0.05)	0.03* (0.02)	0.13*** (0.03)
Fear of Failure	0.26** (0.11)	0.06 (0.04)	0.08 (0.07)
Perceived Good Opportunities	0.25*** (0.09)	0.06 (0.04)	0.01 (0.06)
Fear of Failure x Perceived Good Opportunities	−5.04* (2.55)	−5.13*** (1.10)	−7.14*** (1.80)
CountryBelgium	−2.95 (2.01)	−4.34*** (0.84)	−8.09*** (1.37)
CountryCanada	−3.30* (1.97)	−1.48* (0.85)	−6.02*** (1.38)
CountryChile	17.25*** (2.01)	0.65 (0.84)	−7.07*** (1.36)
CountryCzech Republic	0.28 (3.30)	−0.38 (1.45)	−5.57** (2.36)
CountryDenmark	−4.05** (1.93)	−4.49*** (0.80)	−4.53*** (1.31)
CountryFinland	−5.13*** (1.85)	−4.13*** (0.78)	−2.39* (1.26)
CountryFrance	0.96 (2.81)	−3.92*** (1.16)	−8.59*** (1.89)
CountryGermany	−3.78* (1.99)	−3.81*** (0.82)	−7.02*** (1.34)
CountryGreece	1.14 (2.18)	−3.36*** (0.86)	−1.12 (1.40)
CountryHungary	0.24 (2.16)	−2.60*** (0.90)	−6.32*** (1.47)
CountryIceland	4.31** (1.78)	0.05 (0.75)	−2.99** (1.23)
CountryIreland	−3.35** (1.69)	−2.41*** (0.70)	−4.26*** (1.13)
CountryIsrael	4.27** (2.06)	−3.90*** (0.83)	−7.73*** (1.40)
CountryItaly	−2.33 (1.96)	−4.46*** (0.80)	−6.72*** (1.30)
CountryJapan	−3.50 (3.17)	−4.49*** (1.33)	−2.58 (2.17)
CountryKorea, Republic Of	6.46*** (2.31)	−2.71*** (0.95)	−0.09 (1.55)
CountryMexico	10.17*** (2.05)	1.32 (0.84)	−9.54*** (1.36)
CountryNetherlands	−5.28*** (1.93)	−4.34*** (0.81)	−4.75*** (1.32)
CountryNew Zealand	−0.44 (2.06)	1.18 (0.83)	−1.97 (1.36)
CountryNorway	−3.19* (1.86)	−2.77*** (0.79)	−3.59*** (1.28)
CountryPoland	7.33** (3.22)	−3.50** (1.41)	−4.15* (2.29)
CountryPortugal	−3.51 (2.01)	−4.95*** (0.84)	−6.72*** (1.36)

Table 5: Linear Regression Estimates omitting Country-specific Effects

	<i>Dependent variable:</i>		
	Entrepreneurial Intention	Nascent Entrepreneur	Entrepreneurship Owner
	(1)	(2)	(3)
Way of Teaching	-5.66*** (1.04)	0.15 (0.37)	0.84 (0.56)
Perceived Capabilities	0.31*** (0.04)	0.13*** (0.01)	0.12*** (0.02)
Fear of Failure	0.13 (0.14)	0.02 (0.05)	0.22*** (0.07)
Perceived Good Opportunities	0.24* (0.13)	0.06 (0.04)	0.10 (0.06)
Fear of Failure x Perceived Good Opportunities	-0.005 (0.004)	-0.001 (0.001)	-0.004* (0.002)
Constant	1.94 (5.55)	-3.27* (1.83)	-7.01** (2.76)
Observations	172	192	191
R ²	0.40	0.42	0.20
Adjusted R ²	0.39	0.40	0.18
Residual Std. Error	5.16 (df = 166)	1.94 (df = 186)	2.92 (df = 185)
F Statistic	22.43*** (df = 5; 166)	26.63*** (df = 5; 186)	9.10*** (df = 5; 185)

Note:

*p<0.1; **p<0.05; ***p<0.01

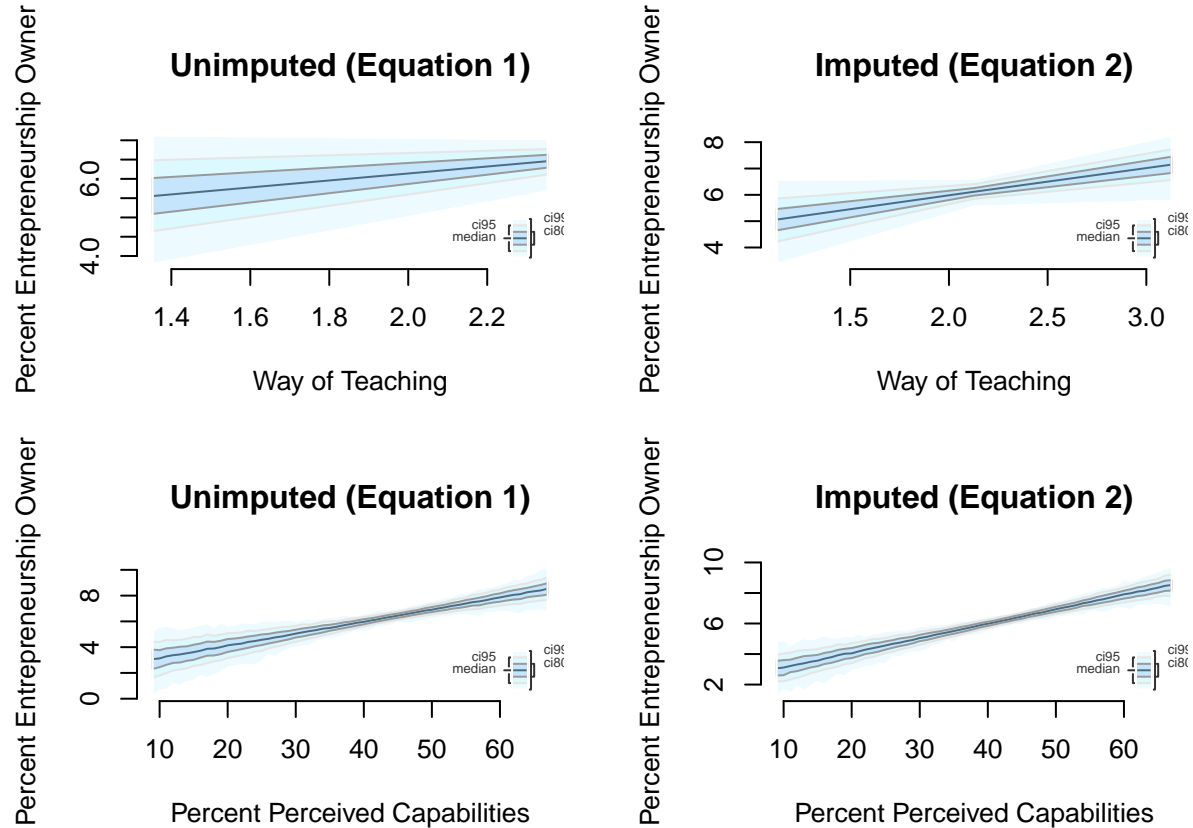


Table 6: Least Squares Regression for Continuous Dependent Variables

	<i>Dependent variable:</i>	
	Entrepreneurship Owner	
	(1)	(2)
Way of Teaching	0.94* (0.52)	1.03*** (0.40)
Perceived Capabilities	0.09*** (0.02)	0.09*** (0.02)
Fear of Failure	0.25*** (0.07)	0.21*** (0.04)
Perceived Good Opportunities	0.14** (0.07)	0.12*** (0.05)
Fear of Failure x Perceived Good Opportunities	-0.005** (0.002)	-0.004*** (0.001)
Constant	-7.67*** (2.78)	-6.82*** (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

As we can see from the regression output table, using the Amelia program to generate additional data expanded the observations analyzed from 172 to 318, and increased the statistical significance of our explanatory variable. In addition, the positive slope of this equation is consistent with what we expected after our analysis of existing literature on this topic. In addition, this process reduced the variance in our simulation results, suggesting it improved the accuracy of this model. Despite this, we have several concerns regarding the reliability of existing GEM data (which will be discussed later in this paper), which makes us question the effectiveness of using this data to generate additional observations.

Perceived opportunities

