Data Article

Survey data on entrepreneurs' subjective plan and perceptions of the likelihood of success

CrossMark

Quan Hoang Vuong

Centre Emile Bernheim, Solvay Brussels School of Economics and Management, Université Libre de Bruxelles,
50 Ave. F.D. Roosevelt, Brussels B-1050, Belgium

A R T I C L E   I N F O

A B S T R A C T

Entrepreneurship is an important economic process in both developed and developing worlds. Nonetheless, many of its concepts appear to be difficult to 'operationalize' due to lack of empirical data; and this is particularly true with emerging economy. The data set described in this paper is available in Mendeley Data's "Vietnamese entrepreneurs' decisiveness and perceptions of the likelihood of success/continuity, Vuong (2015) [1]" http://dx.doi.org/10.17632/kbrtrf6hh4.2; and can enable the modeling after useful discrete data models such as BCL.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Specifications Table

Subject area	*Economics*
More specific subject area	*Business Economics/Entrepreneurship*
Type of data	*Table, text file, graph*
How data was acquired	*Survey*
Data format	*Raw, filtered, and partially analyzed*
Experimental factors	*Raw data obtained from direct survey on participants of seminars, conferences and meetings. Incomplete data sheets are eliminated.*

E-mail address: qvuong@ulb.ac.be

Experimental features	*The experiment focuses on perceptions and subjective understanding of prospective and extant entrepreneurs in Vietnam.*
Data source location	*Hanoi, Ho Chi Minh City, Buon Ma Thuot, Da Nang of Vietnam*
Data accessibility	*Datasets are provided with this article.* Mendeley Data, v2, http://dx.doi.org/ 10.17632/kbrtrf6hh4.2

Value of the data

- The data offer an opportunity to measure the decisiveness and preparedness of an entrepreneur given various conditions that characterize an emerging market.
- Information and deeper insights that might be obtained through discrete data analysis can help predict behaviors of entrepreneurs in typical situations, and formulate policy responses if the government wishes to improve the business/economic environment.
- Important aspects of entrepreneurship such as creativity/innovation, previous professional experience, personal perceptions of socio-cultural values, and the like can be researched and later 'operationalized'.
- The data reflect the transition of the emerging market economy of Vietnam.

1. Data

The data set contains 3071 records obtained from a nationwide survey of perceptions, intentions and assessments from entrepreneurs, existing and prospective, about the socio-economic conditions, values of their previous employment, need of government- and society-supported entrepreneurship-enabling programs. The data also provide subjective evaluation of the likelihood of success or continuity of entrepreneurs' project given certain environmental conditions. The following discrete (categorical) variables are measured in the survey:

Coded name	Explanation	Values
X1.job	Previous job (or the most important)	Human resources (hmr); sales/marketing (salesm), production/operations (pom); finance/accounting (finance); administrative or other departments (admin); no significant job experience (none)
X2.starthis	Entrepreneurial status	Running, dropped, notyet
X3.econdf	Assessment on whether current socio-economic conditions are favorable.	Favorable (fav); somewhat favorable (somewhat); and, unfavorable (unfav)
X4.suppval1	Influence of government or society-funded supports on decision?	High; avg; low
X5.suppval2	Influence of government or society-funded entrepreneurship programs on success?	High; possible; none
X6.infpeople	Influence from family, friends, colleagues.	Strong; somewhat; light; none
X7.tforstart	How much time for this entrepreneurial attempt?	Less than 12 months (less12); 12–24 months (b1224); and, until success (g24)

X8.tforrest	Extra time an entrepreneur gives himself/herself when the project goes bad?	Less than 12 months (less12); 12–24 months (b1224); and, until success (g24)
X9.member	Participation in professional societies, entrepreneurs forums, online networks?	As many as possible (all); in case forums/associations but selective (some); and, none.
X10.failurel	Learning from others' failures?	Careful study (a); exploring few noteworthy cases (b); and, no need (c).
X11.plan	Having a business plan?	A good one (good); a basic plan (basic); in process of making one (inprocess); and, having no need right now (noneed)
X12.resconst	Most serious constraint?	Finance; product; relationship; management skills; technical; others.
X13.mres	Resources for startup	Relatives/friends; investors; gov (donors and government); saving; and, none.
X14.inno	Self-evaluation of creativeness of product/services/business model?	Creative (a), somewhat creative (b), hopefully (c); not at all (d)
X15.diff	Self-evaluation of differentiation of product/services/business model?	Strong (a); Significant differentiation (b); Somewhat different (c); and, not at all (d)
X16.tot1strev	Time to the first dollar (revenue)?	Currently generating revenues (a); in 6 months (b); 6–12 months (c); uncertain (d).
X17. keystrength	Most decisive strength for success	Sales skills (a); good products/services (b); efficient business models (c); accurate prediction of market trend (d); hard working (e); and, patience/tenacity (f).
X18.startplan	When deciding to start own business?	Operating (a); soon (b); within 12 months (c); when conditions are favorable (d); perhaps never (e).
X19.msponge	Entrepreneurial efforts transform ways of thinking, acting and beliefs?	Strong; some aspects; negligible; none
X20. bplantime	When a strategy is needed?	When project begins (a); having revenue (b); firm growing (c); unnecessary (d)
X21.restart	If this turns out a failure, another entrepreneurial attempt?	Never; when possessing new resources; when having new/better ideas; when market conditions and policies support entrepreneurial plan.
X22.csr	Helping next generations of entrepreneurs?	No (a); Yes if that helps improve sales (b); Yes, unconditionally (c).
X23.chance	Evaluated chance of success?	Certain; high ($>80\%$); med (50–80%); low ($<50\%$).
Sex	Gender	Male/Female

2. Experimental design, materials and methods

The survey was designed to obtain discrete data that can be employed by the multi-category logit models to enable analysis based on baseline-category logits (BCL), which helps provide estimated coefficients for computing probabilities upon events of hypothetical influence. The logic for designing the experiment and thus data set is described as follows. For designing both the survey and prepare the data set and suitable subset, an entrepreneur (among n) is treated as independent and identical. Each data point has outcome in any of J categories for each factor to be investigated. Let $y_{ij}=1$ if entrepreneur i has outcome in category j and
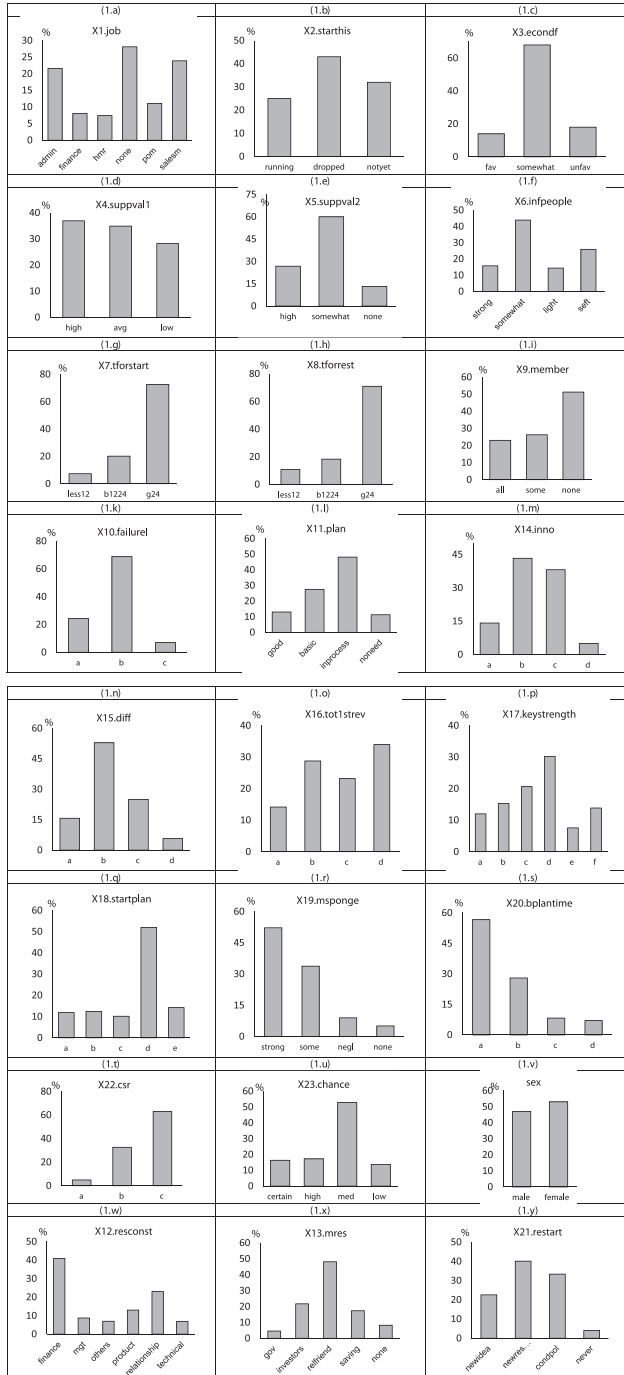
Fig. 1. Some graphs from the raw data.

Table 1
Conditional probabilities of entrepreneurial success based on socio-economic conditions and perceived values of government-funded or society-promoted supporting programs.

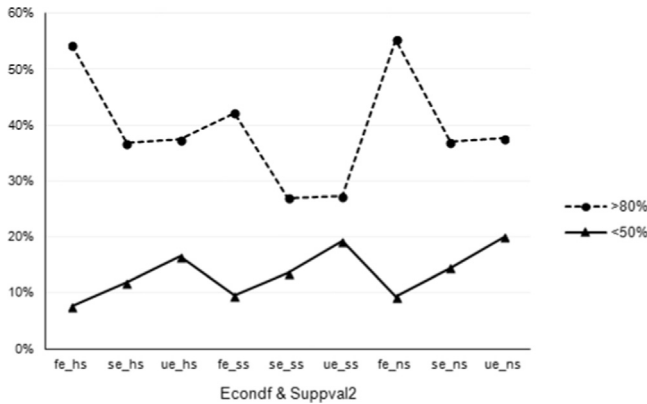| X23.Chance	Certain			High			Med			Low		
X3.econdf\ X5.suppval2	High	s/w	None	High	s/w	None	High	s/w	None	High	s/w	None
Fav	0.335	0.232	0.355	0.208	0.19	0.198	0.381	0.482	0.354	0.076	0.096	0.093
Somewhat	0.17	0.106	0.18	0.198	0.164	0.19	0.514	0.594	0.485	0.118	0.136	0.145
Unfav	0.207	0.132	0.218	0.168	0.141	0.158	0.46	0.535	0.424	0.165	0.192	0.2

Fig. 2. Evaluated chance of success based on economic conditions and values of support.

$y_{ij} = 0$ therwise. Then, $\mathbf{y}_{ij} = (y_{i1}, y_{i2}, ..., y_{ic})$ represents a multinomial trial, with $\sum_j y_{ij} = 1$. As $n_j = \sum_j y_{ij}$ the number of trials having outcome in category j, the design is based on the assumption that $(n_1, n_2, ..., n_c)$ show a multinomial distribution. Let $\pi_j = P(Y_{ij} = 1)$ denote the probability of outcome in category j for each entrepreneur, the multinomial PMF is:

$$p(n_1, n_2, ..., n_c) = \left(\frac{n!}{n_1! n_2! \cdots n_c!}\right) \pi_1^{n_1} \pi_2^{n_2} \cdots \pi_c^{n_c},$$

with:

$$E(n_j) = n\pi_j \; \text{var}(n_j) = n\pi_j(1 - \pi_j) \; \text{cov}(n_j, n_k) = -n\pi_j\pi_k$$

where $\sum_j n_j = n$. As $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$ and $\sum_j \pi_j(\mathbf{x}) = 1$, data are grouped into J categories of Y as multinomial with corresponding sets of probabilities $\{\pi_1(\mathbf{x}), ..., \pi_j(\mathbf{x})\}$. Thus, each response is aligned with a baseline category:

$$\ln \frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} = \alpha_j + \boldsymbol{\beta}_j' \mathbf{x}, \quad j = 1, ..., J-1.$$

BCL analysis simultaneously models the effects of $\mathbf{x}$ on $(J-1)$ logits, which in general vary according to the response paired with the baseline category. The estimating of $(J-1)$ equations employing a given empirical data set would provide for parameters for these logits, as:

$$\ln \frac{\pi_a(\mathbf{x})}{\pi_b(\mathbf{x})} = \ln \frac{\pi_a(\mathbf{x})}{\pi_J(\mathbf{x})} - \ln \frac{\pi_b(\mathbf{x})}{\pi_J(\mathbf{x})}.$$

The empirical data set enables the computing of Pearson-type likelihood ratio test statistics (X^2, G^2) for goodness-of-fit, following a multivariate generalized linear model (GLM) estimations:

$$\mathbf{g}(\boldsymbol{\mu}_i) = \mathbf{X}_i \boldsymbol{\beta},$$

where, $\mathbf{\mu}_i = \mathrm{E}(\mathbf{Y}_i)$, corresponding to $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots)'$; row h of the model matrix $\mathbf{X}_i$ for observation i contains values of independent variables for y_{ih}. For a BCL model, $\mathbf{y}_i = \left(y_{i1}, y_{i2}, \ldots, y_{iJ-1}\right)'$; y_{iJ} is redundant, thus:

$$\mathbf{\mu}_i = (\pi_1(\mathbf{x}_i), \pi_2(\mathbf{x}_i), \ldots, \pi_{J-1}(\mathbf{x}_i))\, g_j(\mathbf{\mu}_i) = \ln\left\{ \mu_{ij} / \left[1 - \left(\mu_1 + \cdots + \mu_{iJ-1} \right) \right] \right\}.$$

Technical details for practical modeling of polytomous logistic models is provided in [2]. Applied analysis can be performed in R (see [3]). Practical uses of survey data can be referred to [4].

Explanation of data subsets filtered for different analysis purposes (from [1]) (Fig. 1).

File name	Filtered with variables:	Frequency distributions table	Appropriate for hypothesis testing of:
x14.15.23	• "X14.inno" • "X15.diff" • "X23. chance"	gt1.1	Creativity and differentiation (entrepreneurs' products/services) influence entrepreneurs' perceptions of the likelihood of success/continuity.
x3.5.23	• "X3. econdf" • "X5. suppval2" • "X23. chance"	gt1.2	Entrepreneurs' assessments on economic conditions and environmental factors (such as support programs) and their impacts on the likelihood of success/continuity.
xsex.7.18	• "sex" • "X7. tforstart" • "X18. startplan"	gt2.1	Gender and strategic intent of timing and duration of efforts by entrepreneurs, in conjunction with their final decision.
x1.3.15.18	• "X1.job" • "X3. econdf" • "X15.diff" • "X18. startplan"	gt2.2	The impact of entrepreneurs' past employment together with self-assessmet of economic conditions, product innovations on the startup decision and likely continuity.
x1.9.15.18	• "X1.job" • "X3. econdf" • "X9. member" • "X18. startplan"	gt2.3	The impact of entrepreneurs' past employment together with self-assessmet of economic conditions, and their networks on the ultimate decisions of starting up and likely continuity.
x11.23.7	• "X11.plan" • "X23. chance" • "X7. tforstart"	gt2.4	Impacts of entrepreneurial planning and perceptions on chance of survival on the timing and likelihood of entrepreneurial undertaking/continuity.

One example of the analysis is to compute response probabilities from multinomial logits, i.e. $\{\pi_j(\mathbf{x})\}$, using $\pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \mathbf{\beta}_j'\mathbf{x})}{1 + \sum_{h=1}^{J-1} \exp(\alpha_h + \mathbf{\beta}_h'\mathbf{x})}$; with $\sum_j \pi_j(\mathbf{x}) = 1$; $\alpha_J = 0$ and $\mathbf{\beta}_J = 0$. An empirical distribution is provided in Table 1.

Picking two different trends, the contrast shown by the empirical data becomes apparent in Fig. 2, suggesting that, if a government aims to promote entrepreneurship, it is better to improve general socio-economic conditions.

Acknowledgments

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.01.034.

References

[1] Q.H. Vuong, Vietnamese entrepreneurs' decisiveness and perceptions of the likelihood of success/continuity, Mendeley Data, 2015, v2 http://dx.doi.org/10.17632/kbrtrf6hh4.2.
[2] A. Agresti, Categorical Data Analysis, Wiley, Hoboken, New Jersey (2002) 774.
[3] PennState Science, Analysis of Discrete Data (STAT 504). ⟨https://onlinecourses.science.psu.edu/stat504/⟩, (accessed 16.10.15).
[4] Q.H. Vuong, Be rich or don't be sick: estimating Vietnamese patients' risk of falling into destitution, SpringerPlus 5 (2015) 529. http://dx.doi.org/10.1186/s40064-015-1279-x.