TI 2006-103/3

Tinbergen Institute Discussion Paper

The Entrepreneurial Ladder and its Determinants

Peter van der Zwan[a]

Roy Thurik[a,c,d]

Isabel Grilo[b]

[a] *CASBEC, Erasmus School of Economics, Erasmus Universiteit Rotterdam, and EIM, Zoetermeer;*

[b] *European Commission, GREMARS, and CORE;*

[c] *Max Planck Institute of Economics, Jena, and Vrije Universiteit Amsterdam;*

[d] *Tinbergen Institute.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

The entrepreneurial ladder and its determinants

Peter van der Zwan[a], Roy Thurik[a,c] and Isabel Grilo[b]

[a] Centre for Advanced Small Business Economics, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands and EIM Business and Policy Research, P.O. Box 7001, 2701 AA Zoetermeer, the Netherlands
pzw@eim.nl

[b] DG Enterprise, European Commission, B-1049, Brussels, Belgium, GREMARS, Université de Lille 3 and CORE, Université Catholique de Louvain.
isabel.grilo@cec.eu.int

[c] Max Planck Institute of Economics, Jena, Germany and Free University Amsterdam
thurik@few.eur.nl

Abstract: We test a new model where the entrepreneurial decision is described as a process of successive engagement levels, i.e., as an entrepreneurial ladder. Five levels are distinguished using nearly 12,000 observations from the 2004 "Flash Eurobarometer survey on Entrepreneurship" covering the 25 European Union member states and the United States. The most surprising of the many results is that perception of lack of financial support is no obstacle for moving to a higher entrepreneurial engagement level whereas perceived administrative complexity is a significant obstacle. We also show that the effect of age on the probability of moving forward in the entrepreneurial process becomes negative after a certain age implying that if entrepreneurial engagements are not taken early enough in life they may well never be taken.

Version: November 2006

Document: orderedlogit pzw rth igr v17.doc 11/21/2006 1:55 PM

JEL-code: H10, J23, L26, M13, R12

Keywords: entrepreneurship, determinants, nascent entrepreneurship, ordered multinomial logit, Europe

Acknowledgement: The authors would like to thank Richard Paap for his useful comments on earlier versions. The views expressed here are those of the authors and should not be attributed to the European Commission. For the first two authors the paper has been written in the framework of the research program SCALES which is carried out by EIM and is financed by the Dutch Ministry of Economic Affairs.

Corresponding author: Roy Thurik, CASBEC, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, thurik@few.eur.nl.

1. Introduction

The theory of occupational choice has dominated the investigations of the entrepreneurship (self-employment) decision (Parker, 2004; Grilo and Thurik, 2005a). It views agents as (expected)-utility maximisers taking an occupational choice decision – to become employees or entrepreneurs – on the grounds of the utility associated with the returns accruing from these two types of activity. Rooted in the work of Knight (1921) this theory sees entrepreneurship as a state which one can adopt or not. This 'static' view has been updated by a more 'dynamic' one acknowledging that setting up a business is a process which consists of several stages (Reynolds, 1997). This view led to a wave of research of the determinants of so-called nascent entrepreneurs (Davidsson, 2006). Nascent entrepreneurs are people who are taking certain steps to become self-employed but are not yet officially established. The work of the Global Entrepreneurship Monitor (GEM) is inspired by this view (Reynolds, et al., 2005).

Grilo and Thurik (2005a and 2006) introduce the concept of engagement levels to discriminate between the various stages of setting up or closing down a business. They apply a simple multinomial logit model to analyze the determinants of the various stages. The engagement levels in the present paper are analyzed in an *ordered* context, in the sense that each level can be associated with an increasing level of involvement in the entrepreneurial process. The idea behind this approach is that entrepreneurship can be described as a process one becomes involved in and where different engagement levels can be distinguished, each having a not necessarily identical set of determinants. (Potential) entrepreneurs climb the entrepreneurial ladder. In the present paper we analyze five of these naturally *ordered* engagement levels. Nearly 12,000 observations are used from the 2004 "Flash Eurobarometer survey on Entrepreneurship" covering all the 25 European Union member states and the United States to analyze whether an *ordered* regression model with five engagement levels gives an adequate description of the entrepreneurial process and to what extent the available covariates are determinants of this process. In other words, we analyze whether these covariates have a significant influence on moving people up the entrepreneurial process.

The contribution of the present paper is that, *first*, while in earlier studies only a multinomial logit model has been used, here we extend this framework to an *ordered* context. Hence, we investigate whether there is a natural ordering of the dependent variable supporting the view of entrepreneurship as a process. *Secondly*, we determine which variables 'drive' (potential) entrepreneurs through this process.

2. Data

In the 2004 "Flash Eurobarometer survey on Entrepreneurship"[1] the following question was used to construct the dependent variable: *"Have you started a business recently or are you taking steps to start one?"* The following options for answering were given:

(1) *"It never came to your mind."*
(2) *"No, but you are thinking about it."*
(2a) *"No, you thought about it or had already taken steps to start a business but gave up."*
(3) *"Yes, you are currently taking steps to start a new business."*
(4) *"Yes, you have started or have taken over a business in the last three years and are still active."*
(5) *"Yes, you started or took over a business more than three years ago and are still active."*
(5a) "*No, you once started a business, but currently you are no longer an entrepreneur."*

Without engagement levels (2a) and (5a) we expect the process to be naturally ordered in terms of involvement in the entrepreneurial process. We will abbreviate the remaining five stages as "Never thought about it", "Thinking about it", "Taking Steps", "Young business" and "Old business", respectively.

[1] http://europa.eu.int/comm/enterprise/enterprise_policy/survey/eurobarometer_intro.htm

Other than demographic variables such as gender (male=1), age, education level (age when finished full time education) and whether parents are self-employed (one or both of the parents are self-employed=1), the set of explanatory variables used includes four perceptions of 'obstacles', a rough measure of risk tolerance, internal and external locus of control and country specific effects. We refer to the usual literature of the determinants of entrepreneurship for justifying the use of these variables (Parker, 2004; Davidsson, 2006; Grilo and Thurik, 2005a, 2005b and 2005c).[2]

The perception variables include the perception by respondents of: lack of available financial support, of complex administrative procedures, of lack of sufficient information on starting an own business, and of an unfavorable economic climate. These variables as well as the risk tolerance variable are captured, respectively, using the question *"Do you strongly agree, agree, disagree or strongly disagree with the following statements?"*:

- *"It is difficult to start one's own business due to a lack of available financial support."*
- *"It is difficult to start one's own business due to the complex administrative procedures."*
- *"It is difficult to obtain sufficient information on how to start a business."*
- *"The current economic climate is not favorable to start one's own business."*
- *"One should not start a business if there is a risk it might fail."*

For the four 'obstacle' statements a dummy variable is constructed which equals 1 in the case of 'strongly agree' or 'agree'. For the 'risk tolerance' statement a dummy variable is constructed which equals 1 if 'disagree' or 'strongly disagree' has been chosen for the fifth statement.

Internal locus of control measures whether an individual believes that (s)he can influence events through own ability, effort or skills. On the other side, external locus of control measures whether an individual believes that external forces determine the outcome. Respondents can choose between five answers on the following question *"When one runs a business, what do you think most determines its success (maximum of two answers)?"*:

- *"The director's personality."*
- *"The general management of the business."*
- *"The overall economy."*
- *"The political context."*
- *"Outside entities."*

The dummy internal success factors equals 1 if one or both of the first two possibilities is mentioned, whereas external success factors equals 1 if one or two of the last three possibilities is mentioned.

Country specific effects are controlled for using country dummies where the US serve as base.

3. Ordered logit model

The ordered logit model builds upon a latent continuous variable, y_i^*, which is modeled using the linear regression $y_i^* = X_i'\beta + \varepsilon_i$, where $i = 1,\ldots,n.$ For example, y_i^* can be thought of as an unobserved willingness to be(come) an entrepreneur. The disturbance terms, ε_i, are uncorrelated and for the ordered logit model it holds that all ε_i follow a logistic distribution with mean zero and variance equal to $\pi^2/3$. X_i is a $k \times 1$ vector of explanatory variables for individual i with corresponding coefficient vector β ($k \times 1$) which is the same across all observations i and engagement levels j.

In contrast with this latent variable we observe the variable Y_i (the engagement level which individual i belongs to) with outcomes y_i, where $y_i = 1,\ldots,J$ and J is the number of engagement levels. Next, y_i^* is related with y_i by means of $J-1$ unobserved thresholds levels $\alpha_1,\ldots,\alpha_{J-1}$:

[2] Following this literature we also apply quadratic terms for age and education next to the linear ones.

$$Y_i = \begin{cases} 1 & \text{if } -\infty < y_i^* \le \alpha_1; \\ j & \text{if } \alpha_{j-1} < y_i^* \le \alpha_j, \text{ for } j = 2,\ldots,J-1; \\ J & \text{if } \alpha_{J-1} < y_i^* \le +\infty. \end{cases}$$

Hence, for $j = 2,\ldots,J-1$, each probability of belonging to engagement level j for individual i is given by $\Pr(Y_i = j) = F(\alpha_j - X_i'\beta) - F(\alpha_{j-1} - X_i'\beta)$ with $F(\cdot)$ the cumulative logistic distribution function. For $j = 1$ we have $\Pr(Y_i = 1) = F(\alpha_1 - X_i'\beta)$ and for $j = J$ this probability equals $1 - F(\alpha_{J-1} - X_i'\beta)$. Note that $X = (X_1,\ldots,X_n)$ does not contain a row of ones for identification purposes.

The above model can be extended to the heteroskedastic case by taking the variance of ε_i to be $E(\varepsilon_i^2) = \frac{1}{3}\pi^2 \exp(z_i'\gamma)^2$ (with z_i a vector of observed variables without constant term) so that $\varepsilon_i / \exp(z_i'\gamma)$ is now a homoskedastic error term. In the remainder we use the notation $\sigma_i = \exp(z_i'\gamma)$. The probability $\Pr(Y_i = j)$ in the heteroskedastic case equals $F\left(\frac{\alpha_j - X_i'\beta}{\sigma_i}\right) - F\left(\frac{\alpha_{j-1} - X_i'\beta}{\sigma_i}\right)$.

4. Model evaluation

The estimation results of both the homoskedastic and heteroskedastic ordered logit model with five engagement levels are shown in Table 1.[3] The magnitude of the coefficients and their significance do not differ much between the two models (only 'education squared' is insignificant in the heteroskedastic formulation). Threshold estimates are of different magnitude in both models but their absolute differences are comparable.

Variables that have a significant influence on the variance of the disturbance term in the heteroskedastic regression are gender (coefficient is 0.108), age (0.018), self-employed parents (0.218), education (-0.007), preference for self-employment (-0.161) (all at a one per cent significance level) and economic climate (0.066) and lack of insufficient info (0.059) (both at five per cent).[4]

Economic interpretation of the heteroskedastic results is somewhat difficult. For instance, one could say that men and older people, ceteris paribus, generate a higher variance of the disturbance term ε_i in the latent regression. In these cases, there is a higher uncertainty in the (latent) value y_i^* and hence, there is more uncertainty about the specific engagement level of the entrepreneurial process an individual belongs to.

Though we have found that the heteroskedastic model is statistically superior to the homoskedastic formulation, we proceed with the interpretation of the homoskedastic model as no important differences are present in the estimation results of the variables and thresholds (apart from 'education squared').

---

[3] We also ran regressions with 1) all engagement levels, 2) only without engagement level (2a) and 3) only without engagement level (5a). In turns out that all diagnostics are in favor of the model we use.

[4] We used a simple likelihood ratio principle to test for the significance of γ in the heteroskedastic specification $\sigma_i = \exp(z_i'\gamma)$. This test statistic, which compares the restricted log-likelihood value (when $\gamma = 0$) with the unrestricted one, is asymptotically χ^2 distributed under the null hypothesis with 7 degrees of freedom (number of restrictions imposed). Note that we did not include a constant in z_i, again due to an identification problem. The resulting value of the test statistic (261.40) is far above the five per cent critical value of a χ^2 distribution with 7 degrees of freedom (14.07) and hence, we reject the null hypothesis of $\gamma = 0$ finding statistically sufficient evidence that the heteroskedastic ordered logit model is preferred to the homoskedastic ordered model.

A crucial assumption underlying the ordered logit model is the 'parallel regression assumption' (same coefficient vector β for each engagement level j). If one includes J engagement levels in the ordered logit model, the test we use (Wald test proposed by Brant, 1990) investigates the equality of the coefficients for all $J-1$ binary logit regressions for k explanatory variables.[5] The coefficient vectors of these $J-1$ logit regressions are denoted as δ_j, $j=1,\ldots, J-1$. The null hypothesis of the Wald test assumes $J-1$ parameter equalities across k variables and hence – as Kim (2003) indicates – we cannot expect this assumption to be true, particularly not in large samples. In our homoskedastic model the 'parallel regression assumption' for *all* variables is violated. One can also check the violation of the 'parallel regression assumption' for each variable separately: only for male, age, age squared, self employed parents and preference for self-employment, the null hypothesis of equal parameter estimates is rejected at one per cent (country dummies are again not considered here). See Table 2 (left hand column). For the variables that do not 'pass the test', it is therefore relevant to look at the results of the binary logit regressions. In Table 2 the estimates of the coefficient vectors δ_j are displayed together with their standard errors as well as marginal effects (not for country dummies).[6] With these marginal effects in mind, one can investigate how impacts of variables change (and the significance of these impacts) with increasing level of involvement.[7] Outcomes are discussed in our section on interpretation.

While testing the 'parallel regression assumption' homoskedasticity is assumed. So, rejection of the 'parallel regression assumption' may be a consequence of not permitting a non-linear function of the latent variable, i.e., a heteroskedastic specification of the error variance. A similar argumentation can be given the other way around: rejecting the homoskedastic specification may be caused by the fact that the 'parallel regression assumption' is not true, i.e., a non-linear specification might be better, while this test is performed under the assumption of equal δ_js.

Allowing for a heteroskedastic specification we test the 'parallel regression assumption' to investigate what the 'real' cause is of rejecting the left side model in Table 1. For each heteroskedastic binary regression we have $\Pr(Y_i = j) = F(x_i'\delta_j^* / \exp(z_i'\gamma_j))$. The estimates of δ_j^* and γ_j^* as well as marginal effects are displayed in Table 3 (without country dummies and constant). The Wald statistic only points at rejection of the 'parallel regression assumption' at a one per cent significance level in the case of preference for self-employment. However, it sometimes gives negative values. The results for gender, age, age squared, self-employed parents and administrative complexities tend to show less spread across the four binary regressions than the results of homoskedastic binary regression given in Table 2. As said above, for these five variables the 'parallel regression assumption' is violated in the homoskedastic case while the coefficients are significant. It is tempting to conclude that rejection of the 'parallel regression assumption' in the homoskedastic model is due to not allowing for a heteroskedastic formulation.[8]

5. Interpretation

Interpretation of the ordered logit model is best done using the log odds ratios $\log(\Pr(Y_i \le j)/\Pr(Y_i > j)) = \alpha_j - X_i'\beta$. So, for each engagement level j, a positive coefficient implies

[5] To illustrate these binary regressions, suppose one has three engagement levels, so $J = 3$. One can now perform two separate binomial logit regressions: $\Pr(Y_i = 1)$ versus $\Pr(Y_i > 1)$ and $\Pr(Y_i \le 2)$ versus $\Pr(Y_i = 3)$. For each binary regression a different coefficient vector is estimated. When these coefficient vectors do not significantly differ from each other, there is no reason to reject the 'parallel regression assumption'.

[6] The computation of the marginal effects is done as follows: for each observation a marginal effect is calculated and the sample averages of these values are displayed in Table 2 for all variables. The p-values of these effects are comparable to p-values of the coefficients of the binary regressions in the same table.

[7] If the 'parallel regression assumption' is not violated for a variable, this does not necessarily imply that the marginal effects in Table 2 are statistically the same across all binary regressions.

[8] Furthermore, we investigated the redundancy of the variables in the heteroskedastic specification (testing $\gamma_j^* = 0$ for each j) with a likelihood ratio test statistic (7 degrees of freedom, 0.05 critical value is 14.07). The four test statistics given in Table 3 (79.42; 69.08; 58.20; 51.22) are all in excess of 14.07. We did the same for the entire model (46 degrees of freedom, 0.05 critical value is 62.83). The four test statistics given in Table 3 (3343.66; 2034.88; 1776.52; 1351.76) are all in excess of 62.83.

that an increase in the corresponding variable, while keeping all other variables equal, leads to a situation where an individual is more likely to move to an engagement level higher than *j* than to stay in *j*.

The estimates of the thresholds show that the first is relatively far away from the second (the confidence intervals do not even overlap). It seems difficult to switch from "Thinking about it" to "Taking steps". Once in the entrepreneurial process, the step from "Taking steps" to "Young business" is relatively easily made. This gap again is smaller than the one from "Young business" to "Old business".[9]

Demographic variables: gender, age, education

Table 1 reveals that the gender coefficient is significantly different from zero: men have a higher probability than women of moving to a higher level of entrepreneurial involvement. Note that for this gender variable the 'parallel regression assumption' has been violated, because of a different coefficient in each binary regression (see Table 2). Furthermore we see in Table 2 that the effect of gender on the probability of being in engagement level *j+1* versus *j* decreases as *j* increases. So, the effect of gender becomes weaker (it plays a less important role) when higher levels of engagement are attained.

As can be seen from Table 1 age and education are significantly present in the ordered regression. Because of the violation of the 'parallel regression assumption' for the age variable we take a further look at Table 2. Taking into account the squared term we can calculate the turning points at which the effect of age becomes negative for each binary regression. It turns out that these turning points vary between 36 years old for the switch from "Never thought about it" to higher levels of involvement and 51 years old in the last binary regression which confronts any level of engagement below having a business for at least 3 years versus the highest involvement level of being an owner for at least 3 years. These turning points increase steadily as the switch portrayed in the binary regression corresponds to higher levels of entrepreneurial involvement.[10] These results seem to suggest that the 'jump' into any form of entrepreneurial involvement, even the mildest "Thinking about it", is more likely to be made until the mid-thirties with age playing against it as one gets older than that. Without making a case of the precision of this specific age, what this result implies as a message for those who design measures or incentives to help people consider an entrepreneurial carrier, is that the chances of success in triggering such a change of mind decrease after a certain age. In the same vain, using the information conveyed by the turning points implicit in the other binary regressions, every move towards higher levels of entrepreneurial engagement is less likely after a certain age.[11] These results, eventually complemented by additional research, are useful for policy makers in determining target groups depending on the type of measures envisaged to prompt an entrepreneurial response from the population.

For education, on the other hand, the 'parallel regression assumption' has not been violated (the coefficient stays the same across all engagement levels). Furthermore, despite the negative sign of 'education squared' in Table 1 the effect of education remains positive in the relevant range.[12]

Self-employment preference and self-employed parents

Preference for self-employment is significantly present in the ordered regression. This coefficient does not change as one becomes more active in the entrepreneurial world. The marginal effect of this variable, however, decreases heavily in moving forward in the entrepreneurial process, while this variable seems to be very important in the switching behavior as can be seen from the large marginal effects across all binary regressions.

---

[9] These results support the use of the influential TEA (Total Entrepreneurial Activity) measure of GEM where nascent and young entrepreneurs are taken together (Reynolds et al., 2005).

[10] For each binomial regression in Table 2 the turning point where the effect of age becomes negative is 36, 46, 48 and 51 years old. These numbers are similar to those obtained in the heteroskedastic binary regressions, except that the turning point of any level of engagement below having a business for at least 3 years versus the highest involvement level of being an owner for at least 3 years becomes 50 years in stead of 51.

[11] Reynolds (1997) using the concept of "nascent entrepreneurs" (those reporting two or more firm gestation behaviours) finds that age is the dominant factor affecting decisions to start a new firm and that this effect is non-monotonic attaining its peak for the age class 25 to 34.

[12] The turning point for education resulting from the coefficients in Table 1 takes the value of 47 for the variable "age when finished full time education".

Having self-employed parents also significantly increases the probability of moving to higher engagement levels, as the (large) significant marginal effects in Table 2 reveal.

Obstacle variables

The perception of lack of financial support does not affect the probability of being in any of the stages of the entrepreneurial process. It does not seem to discourage respondents in setting up a business and becoming entrepreneur. The same holds true for the lack of sufficient information. Also, the fact of perceiving an unfavorable economic climate does not play a role in switching through the whole entrepreneurial system, although in the last two binary regressions concerning levels of high involvement, this variable *does* have a significant effect.

The fact that a respondent perceives it to be difficult to start a business due to complex administrative procedures has a negative impact on the probability of being in the more 'active' levels of entrepreneurship (see the significant negative coefficient estimate in Table 1 and the significant negative marginal effects in Table 2). Furthermore, if one is more risk tolerant, one is more likely to move to a higher engagement level in the entrepreneurial system than staying in the present engagement level.

Internal and external locus of control

Finally, internal and external success factors do not seem to be relevant in the context of the present setup. Hence, the fact that an individual believes that he or she can influence events through his/her own ability or skills does not have a significant influence of being in one of the five stages of the entrepreneurial process. The same can be concluded for the acknowledgement that external factors influence events.

Country dummies

Parameter estimates of the country dummies are insignificant in case of Denmark, Greece, Netherlands, United Kingdom, Latvia, Poland and Slovenia (at the ten per cent significance level). Furthermore, the coefficients are largest for Belgium, Spain, France, Portugal and Malta (all negative) and Estonia and Slovakia (both positive).

6. Conclusion

We start from the assumption that the decision to become entrepreneur should be modeled as a process rather than as a binary choice. We discriminate between five stages of entrepreneurship (engagement levels). These stages are successive so that 'climbing the entrepreneurial ladder' becomes the obvious metaphor. For each stage, 2004 survey data are available at the individual level for all the 25 EU member states and the US. We analyze these engagement levels using an ordered logit model to investigate the influence of various explanatory variables on moving through the various stages of the process, i.e., on climbing the ladder.

The estimation results of the ordered logit threshold levels reveal that it is difficult to switch from "Thinking about starting a business" to "Taking steps to start a business". Once in the entrepreneurial process, the step from "Taking steps" to "Having a young business" is made more easily. This gap is smaller than the one from "Having a young business" to "Having an old business".

We have shown that the effects of gender and education are positive and significant while those of age are positive up to a certain age, at which point they turn negative. Moreover, on the basis of a set of binary regressions it is shown that the turning point at which the effect of age turns negative increases with higher levels of entrepreneurial involvement. Men move more easily through the process than women while the effect of this variable decreases with the level of entrepreneurial involvement. Furthermore, the better educated people move more easily through the process. Also, if one has a preference for self-employment, one is more likely to move to a higher engagement level than to stay in the current one. While the perception of lack of financial support, of insufficient information and of an unfavorable economic climate do not have a significant impact (this last variable has significant effects in the switching from "Taking steps" to "Young business" and from "Young business" to "Old business"), a respondent's perception that it is difficult to start a business due to complex administrative

procedures has a negative impact on switching to higher engagement levels. Besides, more risk tolerant people find it easier to move upward through the various stages than people who are less risk tolerant.[13]

In this conclusion we want to stress the policy implications of two findings. First, we found that beyond the age of 36 years the probability of at least thinking about embracing an entrepreneurial carrier decreases. Together with the phenomenon of the aging European societies, this finding gives a sense of urgency to policies aimed at turning potential entrepreneurs into active ones. Second, our finding that administrative complexities have a negative effect on the probability of moving forward in the entrepreneurial process lends support to the many public efforts to cut red tape and adopt better regulation approaches.

7. References

Brant, R. (1990), 'Assessing proportionality in the proportional odds model for ordinal logistic regression', *Biometrics*, 46 (4), 1171-1178.

Davidsson, P. (2006), 'Nascent entrepreneurship: empirical studies and developments', *Foundations and Trends in Entrepreneurship Research,* 2 (1), 1-76.

Grilo, I. and J.M. Irigoyen (2006), 'Entrepreneurship in the EU: to wish and not to be', *Small Business Economics*, 26 (4), 305-318.

Grilo, I. and A.R. Thurik (2005a), 'Determinants of entrepreneurial engagement levels in Europe and the US', Max Planck Discussion Paper on Entrepreneurship, Growth and Public Policy nr. 2505, Jena, Germany: Max Planck Institute of Economics.

Grilo, I. and A.R. Thurik (2005b), 'Latent and actual entrepreneurship in Europe and the US: some recent developments', *International Entrepreneurship and Management Journal*, 1, 441-459.

Grilo, I. and A.R. Thurik (2005c), 'Entrepreneurial engagement levels in the European Union', *International Journal of Entrepreneurship Education*, 3 (2), 143-168.

Kim, J.-H. (2003), 'Assessing practical significance of the proportional odds assumption', *Statistics and Probability Letters*, 65 (3), 233–239.

Knight, Frank (1921), *Risk, Uncertainty, and Profit*, Boston: Houghton Mifflin Company.

Parker, S.C. (2004), *The Economics of Self-Employment and Entrepreneurship*, Cambridge: Cambridge University Press.

Reynolds, P. (1997), 'Who starts new firms? – Preliminary explorations of firms-in-gestation', *Small Business Economics*, 9 (5), 449–462.

Reynolds, P.D., N. Bosma, E. Autio, S. Hunt, N. de Bono, I. Servais, P. Lopez-Garcia and N. Chin (2005), 'Global entrepreneurship monitor: data collection design and implementation 1998-2003', *Small Business Economics* 24(3), 205-231.

[13] The absence of a significant impact of the perception of lack of financial support as well as the unambiguous influences of the perception of administrative complexities, preference for self-employment and risk tolerance are in line with findings in earlier studies using different probit models but also based on the "Flash Eurobarometer survey on Entrepreneurship" data sets of different years (Grilo and Irigoyen, 2006; Grilo and Thurik, 2005b).

Table 1. Estimation results ordered logit model (estimates of coefficient vector β and threshold levels with corresponding standard errors).

	Homoskedastic			Heteroskedastic		
	coeff.		std.err.	coeff.		std.err.
Gender	0.547	***	0.041	0.806	***	0.096
Age	0.134	***	0.007	0.317	***	0.030
(Age/100) squared	-16.864	***	0.841	-44.869	***	4.300
Education	0.068	***	0.013	0.115	***	0.030
(Education/100) squared	-7.264	***	2.534	-9.264		5.983
Self-employed parents	0.398	***	0.046	0.464	***	0.104
Lack financial support	-0.019		0.053	-0.100		0.097
Administrative complex.	-0.192	***	0.047	-0.306	***	0.088
Insufficient info	0.052		0.044	0.008		0.087
Risk tolerance	0.169	***	0.043	0.254	***	0.081
Economic climate	0.029		0.046	-0.056		0.090
Preference self-employment	1.756	***	0.045	3.539	***	0.251
Internal success factors	-0.030		0.049	-0.062		0.091
External success factors	-0.064		0.055	-0.100		0.105
Belgium	-0.725	***	0.133	-1.403	***	0.259
Denmark	-0.029		0.157	-0.218		0.303
Germany	0.216	*	0.117	0.260		0.228
Greece	0.172		0.112	0.194		0.212
Spain	-0.918	***	0.129	-1.846	***	0.259
France	-0.874	***	0.129	-1.680	***	0.270
Ireland	-0.491	***	0.144	-0.940	***	0.265
Italy	-0.546	***	0.116	-1.176	***	0.234
Luxembourg	-0.572	***	0.156	-1.217	***	0.284
Netherlands	0.157		0.124	0.308		0.244
Austria	0.319	**	0.160	0.360		0.301
Portugal	-0.584	***	0.124	-1.383	***	0.244
Finland	0.369	**	0.154	0.562	*	0.288
United Kingdom	-0.023		0.122	-0.002		0.232
Czech Republic	0.334	***	0.125	0.634	**	0.247
Estonia	0.700	***	0.148	1.114	***	0.306
Cyprus	-0.394	***	0.147	-0.861	***	0.258
Latvia	0.009		0.140	-0.057		0.267
Lithuania	0.339	**	0.139	0.559	**	0.268
Hungary	0.237	*	0.128	0.207		0.237
Malta	-0.620	***	0.171	-1.182	***	0.318
Poland	0.015		0.118	-0.070		0.207
Sweden	-0.359	**	0.156	-0.787	***	0.293
Slovakia	0.746	***	0.140	1.373	***	0.297
Slovenia	0.230		0.142	0.373		0.266
Threshold 1	4.876	***	0.239	9.302	***	0.711
Threshold 2	6.492	***	0.243	12.469	***	0.913
Threshold 3	6.855	***	0.244	13.220	***	0.967
Threshold 4	7.355	***	0.245	14.309	***	1.046
Number of observations	11751			11751		
Log-likelihood	-10927.83			-10666.40		
LR statistic	3349.30 (χ^2, 39 df.)			3872.16 (χ^2, 46 df.)		
Akaike inform. crit.	1.867			1.824		
Bayesian inform. crit.	1.894			1.855		
McFadden R^2	0.133			0.154		

***: significant at 0.01; **: at 0.05; *: at 0.10.

Table 2. Results from four homoskedastic binary logit regressions (estimates of coefficient vectors δ_j, together with average marginal effects).

	Binary regression							
	(1) vs. >(1)		<=(2) vs. >(2)		<=(3) vs. >(3)		<=(4) vs. (5)	
	coeff.	effect	coeff.	effect	coeff.	effect	coeff.	effect
Gender[###, ^^^]	0.509***	0.091	0.753***	0.077	0.819***	0.067	0.853***	0.049
Age[###, ^^^]	0.104***	0.018	0.241***	0.025	0.306***	0.025	0.328***	0.019
(Age/100) squared[###, ^^^]	-14.498***	-2.548	-26.190***	-2.687	-31.671***	-2.615	-32.273***	-1.908
Education[^^^]	0.068***	0.012	0.068***	0.007	0.079***	0.007	0.064**	0.004
(Education/100) squared[^^^]	-6.368**	-1.119	-8.934**	-0.917	-13.017***	-1.075	-11.827**	-0.699
Self-employed parents[###, ^^^]	0.340***	0.061	0.608***	0.067	0.685***	0.061	0.684***	0.044
Lack financial support	-0.003	-0.001	-0.069	-0.007	-0.023	-0.002	-0.063	-0.004
Administr. complex.[##,^^^]	-0.143***	-0.025	-0.283***	-0.030	-0.338***	-0.029	-0.270***	-0.017
Insufficient info[##]	0.042	0.007	0.162**	0.017	0.114	0.009	-0.005	0.000
Risk tolerance[^^^]	0.167***	0.029	0.200***	0.021	0.246***	0.020	0.200**	0.012
Economic climate[##]	-0.003	-0.001	0.080	0.008	0.184**	0.015	0.235***	0.013
Preference for self-employment[###,^^^]	1.783***	0.348	1.758***	0.178	1.605***	0.130	1.654***	0.093
Internal success factors	-0.076	-0.013	0.061	0.006	0.082	0.007	0.076	0.005
External success factors	-0.101*	-0.018	0.042	0.004	-0.001	0.000	0.040	0.002
Number of observations	11751		11751		11751		11751	
Log-likelihood	-6183.25		-3930.88		-3269.68		-2466.62	
LR statistic (χ^2, 39 df.)	3264.23		1965.79		1718.32		1300.54	
Akaike inform. crit.	1.059		0.676		0.563		0.427	
Bayesian inform. crit.	1.084		0.701		0.588		0.452	
McFadden R^2	0.209		0.200		0.208		0.209	

***: coefficient and marginal effect significant at 0.01; **: at 0.05; *: at 0.10.

###: 'parallel regression assumption' violated for this variable at 0.01; ##: at 0.05; #: at 0.10.

^^^: significant at 0.01 in homoskedastic logit regression (see Table 1).

Table 3. Results from four heteroskedastic binary logit regressions (estimates of coefficient vectors δ_j^* and γ_j^*, together with average marginal effects).

	Binary regression							
	(1) vs. >(1)		<=(2) vs. >(2)		<=(3) vs. >(3)		<=(4) vs. (5)	
	coeff.	effect	coeff.	effect	coeff.	effect	coeff.	effect
Gender^^^	0.413***	0.086	0.409***	0.074	0.518***	0.064	0.499***	0.047
Age^^^	0.088***	0.018	0.168***	0.023	0.235***	0.024	0.248***	0.018
(Age/100) squared^^^	-12.247***	-2.426	-18.315***	-2.486	-24.342***	-2.470	-24.745***	-1.784
Education^^^	0.046***	0.009	0.046***	0.007	0.064***	0.006	0.040*	0.003
(Education/100) squared	-3.408	-0.675	-6.183**	-0.839	-10.331***	-1.048	-7.918*	-0.571
Self-employed parents^^^	0.229***	0.056	0.169**	0.060	0.168*	0.056	0.068	0.041
Lack financial support	-0.028	-0.006	-0.068	-0.009	-0.027	-0.003	-0.071	-0.005
Administr. complex.^^^	-0.100**	-0.020	-0.186***	-0.026	-0.259***	-0.027	-0.231***	-0.017
Insufficient info	0.000	0.009	-0.030	0.015	-0.066	0.008	-0.051	0.000
Risk tolerance^^^	0.121***	0.024	0.135***	0.018	0.174***	0.018	0.157**	0.011
Economic climate	-0.008	0.002	0.001	0.010	0.043	0.015	0.003	0.014
Preference for self-employment^^^	2.002***	0.356	2.958***	0.183	2.822***	0.133	3.867***	0.095
Internal success factors	-0.058	-0.012	0.056	0.008	0.085	0.009	0.088	0.006
External success factors	-0.096*	-0.019	0.027	0.004	-0.017	-0.002	0.013	0.001
Gender	0.034		0.125***		0.089*		0.106*	
Age	0.000		0.000		0.001		0.002	
Education	0.004		0.002		-0.001		0.002	
Self-employed parents	0.169***		0.260***		0.282***		0.300***	
Insufficient info	0.124***		0.139***		0.117**		0.033	
Economic climate	0.053		0.070		0.084		0.135**	
Preference for self-employment	-0.573***		-0.746***		-0.646***		-0.808***	
Number of observations	11751		11751		11751		11751	
Log-likelihood	-6143.54		-3896.34		-3240.58		-2441.01	
LR statistic (χ^2, 7 df.)	79.42		69.08		58.20		51.22	
LR statistic (χ^2, 46 df.)	3343.66		2034.88		1776.52		1351.76	
Akaike inform. crit.	1.054		0.671		0.560		0.423	
Bayesian inform. crit.	1.083		0.701		0.589		0.453	
McFadden R^2	0.214		0.207		0.215		0.217	

***: coefficient and marginal effect significant at 0.01; **: at 0.05; *: at 0.10.

^^^: significant at 0.01 in heterokedastic logit regression (see Table 1).