H200719

Modelling latent and actual entrepreneurship

Isabel Grilo
Peter van der Zwan
Roy Thurik

Zoetermeer, December 2007

SCALES
SCientific AnaLysis of Entrepreneurship and SMEs

This report is published under the SCALES-initiative (SCientific AnaLysis of Entrepreneurship and SMEs), as part of the 'SMEs and Entrepreneurship programme' financed by the Netherlands Ministry of Economic Affairs.

Most recent EIM reports and much more on SMEs and Entrepreneurship can be found at: www.eim.nl/smes-and-entrepreneurship.

Modelling latent and actual entrepreneurship

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

[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
vanderzwan@few.eur.nl; thurik@few.eur.nl;

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

[c] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands
paap@few.eur.nl

[d] Max Planck Institute of Economics, Jena, Germany and Free University Amsterdam

Abstract: The determinants of latent (i.e., desired) and actual entrepreneurship are analysed in two ways with nearly 8,000 observations from the 2004 "Flash Eurobarometer survey on Entrepreneurship" covering the 25 European Union member states and the United States. Both methods lead to new and extensive insights in the interrelation of both concepts. First, latent and actual entrepreneurship are investigated simultaneously in a bivariate probit setting. The perception of lack of financial support, the perception of administrative complexities, and the perception of lack of sufficient information do not have significant direct impacts on latent entrepreneurship. This points at indirect effects of these variables on latent entrepreneurship via actual entrepreneurship. Second, four groups of individuals are distinguished, based on their involvement in both measures of entrepreneurship. The analysis enables us for example to discuss the determinants of 'necessity entrepreneurship'. Results show that the perception of administrative complexities is a significant obstacle in setting up a business, irrespective of the declared preference for self-employment, while the perception of financial constraints does not have a significant influence. Also, necessity entrepreneurs are characterized by a relatively low education level compared to those who are neither latent nor actual entrepreneurs. Each model has its own merits. The multinomial model enables researchers to perform group-wise analyses, while the bivariate probit model makes is possible to take into account the importance of latent entrepreneurship without explicitly including latent entrepreneurship in the set of explanatory variables.

Version: December 2007

Document: actuallatent pzw igr rth v4 1/16/2008 9:59 AM

JEL-code: H10, J23, L26, M13

Keywords: entrepreneurship, latent entrepreneurship, bivariate probit, 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 last 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 decision to become self-employed is widely investigated, while only few studies have addressed the underlying desire to become self-employed, i.e., latent entrepreneurship. The present study analyses the explanation of this entrepreneurial spirit in the European Member states and the United States.

In Blanchflower *et al*. (2001) and Grilo and Irigoyen (2006), a declared preference of being self-employed over being an employee is used as an approximation of latent entrepreneurship. These studies find that latent entrepreneurship depends positively on gender and negatively on age while the number of years of schooling does not have a significant effect. In addition, risk tolerant people are more likely to reveal a preference for self-employment (see also Grilo and Thurik, 2005a). Masuda (2006) focuses on Japan and concludes that unemployment rates and total cash earnings (as a substitute for demand or income growth) have positive effects on latent entrepreneurship. Two measures of latent entrepreneurship are used in the particular study. One refers to the amount of people of the work force wishing to change their current jobs in favour of self-employment. The other measures the number of working people preparing to become self-employed.

The present study differs in methodology compared to earlier studies on latent entrepreneurship in a few ways. In Blanchflower *et al*. (2001), Grilo and Irigoyen (2006) and Grilo and Thurik (2005a), two independent probit equations are estimated, where one equation explains latent entrepreneurship based on a set of explanatory variables and the other equation explains actual entrepreneurship depending on the same set of explanatory variables and latent entrepreneurship. The present study, however, first focuses on the simultaneous explanation of latent and actual entrepreneurship with emphasising the different points in time of the realisation of both concepts implying that latent entrepreneurship (as used in this study) also depends on the actual self-employment status. In a next phase we apply a multinomial logit model in which joint probabilities of latent and actual entrepreneurship are modelled. We also consider each approach on its own merits. Finally, the set of explanatory variables is extended compared to earlier mentioned studies on latent entrepreneurship.

2. Data

Data from the 2004 *"Flash Eurobarometer survey on Entrepreneurship"*[1] are used. The analysis is therefore concentrated on the individual level in 25 EU member states and the United States.[2]

Two different indicators of entrepreneurship will be used. The first aims at capturing the entrepreneurial drive of respondents ("latent entrepreneurship"). To define this concept the following question is used: *"Suppose you could choose between different kinds of jobs, which one would you prefer: being an employee or being self-employed?"*

It has already been discussed in scholarly work whether this simplified concept of latent entrepreneurship represents entrepreneurial preferences in an adequate way (Blanchflower *et al*., 2001; Grilo and Irigoyen, 2006).[3] Although respondents are confronted with a hypothetical

[1] See http://europa.eu.int/comm/enterprise/enterprise_policy/survey/eurobarometer_intro.htm for more information about this survey.

[2] In 2004 and 2005 the number of countries in the EU was equal to 25. From January 1, 2007 onwards Rumania and Bulgaria are also member states. These countries are therefore not included in the data set.

[3] Grilo and Irigoyen (2006) verbalise this as follows: "Its hypothetical flavor may unleash a value judgment over some attractive attributes associated with self-employment – independence, higher income, opportunity of tax evasion – without

setting of being self-employed without considering negative consequences that would have been experienced in a more realistic setting, we are convinced of the usefulness of the measure as this very same question is asked in each country.

The following question provides the basis of a proxy for actual entrepreneurship. *"As far as your current occupation is concerned, would you say you are self-employed, an employee, a manual worker or would you say that you are without a professional activity?"*

In Table 1 the discrepancy between latent and actual entrepreneurship is put forward. This table shows that entrepreneurial spirit differs largely across countries, from the US, Portugal, Cyprus, Spain and Ireland (above 60 per cent of the respondents) to the Netherlands, Slovakia and Finland (in these countries, there exists a minor drive for setting up a business). In no more than eight countries the "transmission" from latent to actual entrepreneurship exceeds 50 per cent (Finland, Greece, Slovakia and the Netherlands top this ranking). It is this difference between the two measures that gives rise to the present study.

Other than demographic variables such as gender (male = 1; female = 0), 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; otherwise 0), the set of explanatory variables includes four perceptions of "obstacles", a crude 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 2008).

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 unfavourable 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 favourable 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.

Ideally one would have the availability of the perception variables at the moment of implementing the entrepreneurship decision. It is important to keep the following arguments in mind. These perception questions can be interpreted ambivalently by respondents as respondents may think of 1) their own situation, or 2) they may think of the general attitude towards entrepreneurship in a particular country or a region (independent of their own experience). Besides, there is another ambiguity concerning the wording of the particular questions. Because of the "due to" part in each item, basically two issues are addressed in the very same question: the difficulty of starting an own business, and the relationship of the barriers with this difficulty. It may be that the first part of each item dominates the second part of the particular item and therefore respondents may be less focused on the second part and as a result the answers mainly reflect the difficulty of starting an own business. Another problem

considering all the consequences of a more realistic setting."

may arise in case respondents do not find it difficult to start an own business at all. It is unclear in this case how the answers to the questions should be interpreted. Most probably, the answers refer to the general attitude towards entrepreneurship in a country or a region.

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 to the question *"When one runs a business, what do you think determines most 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 without mentioning any of the last three, whereas external success factors equals 1 if one or more of the last three possibilities is mentioned without again mentioning any of the first two.

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

3. Latent and actual entrepreneurship in a bivariate probit setting

3.1. Introduction and motivation

In this section the determinants of latent and actual entrepreneurship are simultaneously analyzed. Our approach differs from the approaches in Grilo and Irigoyen (2006) and Grilo and Thurik (2005a) in at least two ways. Besides the fact that latent entrepreneurship also depends on the actual status of employment, the two equations are estimated simultaneously in addition. That is, we estimate a simultaneous pair of probit equations in which actual entrepreneurship is explained by latent entrepreneurship and explanatory variables and latent entrepreneurship is explained by actual entrepreneurship and a (nearly identical) set of explanatory variables.

We now give a rationale behind the inclusion of actual entrepreneurship as an independent variable in the equation for latent entrepreneurship. Latent entrepreneurship is approximated with a declared preference of individuals for self-employment. Preferably, one would measure the preference for self-employment at the moment of implementing the decision of becoming an entrepreneur. This variable is, however, not available and therefore the past (stated) preference is approximated with a declared preference at the moment of the survey. This declared preference is a mixture of information the respondent has about his/her working environment and of the interpretation of this information by the particular respondent. The declared preference therefore is likely to depend on the actual employment status of the respondent.

This chapter and the next chapter essentially have the same objective, i.e., analyzing the determinants of latent and actual entrepreneurship. This chapter explores a bivariate probit model that includes latent entrepreneurship and other covariates as explanatory variables; the next chapter focuses on a group-specific analysis. In that particular analysis, one is able to estimate the effects of variables on, for example, actual entrepreneurship with taking into account the importance of latent entrepreneurship but without explicitly including latent entrepreneurship in the set of explanatory variables.

3.2. Simultaneous equations framework

Ideally, one would have the following system of two simultaneous latent equations[4]:

$$y_{i1}^* = x_{i1}'\beta_1 + \gamma_1 y_{i2} + \varepsilon_{i1}, \qquad y_{i1} = \begin{cases} 1 & \text{if} \quad y_{i1}^* > 0; \\ 0 & \text{if} \quad y_{i1}^* \le 0. \end{cases}$$

$$y_{i2}^* = x_{i2}'\beta_2 + \gamma_2 y_{i1} + \varepsilon_{i2}, \qquad y_{i2} = \begin{cases} 1 & \text{if} \quad y_{i2}^* > 0; \\ 0 & \text{if} \quad y_{i2}^* \le 0. \end{cases}$$

We now suppose that y_{i1} is equal to actual entrepreneurship (with values 0 and 1) while y_{i2} equals latent entrepreneurship (also values 0 and 1) with corresponding underlying latent variables y_{i1}^* and y_{i2}^*. This system is equivalent to the probit equations $\Pr(y_{i1}=1) = \Phi(x_{i1}'\beta_1 + \gamma_1 y_{i2})$ and $\Pr(y_{i2}=1) = \Phi(x_{i2}'\beta_2 + \gamma_2 y_{i1})$ with $\Phi(\cdot)$ the univariate standard normal distribution function. The model is not identified as is shown in Maddala (1983).[5] Therefore, we now consider an analogous formulation which does not include the actually observed realizations of latent and actual entrepreneurship; instead, the latent counterparts are present in both the left and the right-hand side. This pair of simultaneous equations reads as:

$$y_{i1}^* = x_{i1}'\beta_1 + \gamma_1 y_{i2}^* + \varepsilon_{i1}, \qquad y_{i1} = \begin{cases} 1 & \text{if} \quad y_{i1}^* > 0; \\ 0 & \text{if} \quad y_{i1}^* \le 0. \end{cases}$$

$$y_{i2}^* = x_{i2}'\beta_2 + \gamma_2 y_{i1}^* + \varepsilon_{i2}, \qquad y_{i2} = \begin{cases} 1 & \text{if} \quad y_{i2}^* > 0; \\ 0 & \text{if} \quad y_{i2}^* \le 0. \end{cases}$$

Identification of this model is fulfilled given that x_{i1} contains at least one variable that is not contained in x_{i2} and x_{i2} contains at least one variable that is not contained in x_{i1} (x_{i1} and x_{i2} both have length k).

To find consistent estimates γ_1, γ_2, β_1 and β_2 reduced form equation have to be estimated in which the endogenous variables y_{i1}^* and y_{i2}^* depend on exogenous variables only (only the terms $x_{i1}'\beta_1$ and $x_{i2}'\beta_2$ are included at the right-hand side). Replacing the expression for y_{i2}^* in the equation with y_{i1}^* as dependent variable and doing the same for the expression of y_{i1}^* we arrive at the following expressions:

$$(1-\gamma_1\gamma_2)y_{i1}^* = x_{i1}'\beta_1 + x_{i2}'\gamma_1\beta_2 + (\varepsilon_{i1} + \gamma_1\varepsilon_{i2});$$
$$(1-\gamma_1\gamma_2)y_{i2}^* = x_{i2}'\beta_2 + x_{i1}'\gamma_2\beta_1 + (\varepsilon_{i2} + \gamma_2\varepsilon_{i1}),$$

or

$$y_{i1}^* = \pi_1 X_i + \nu_{i1};$$
$$y_{i2}^* = \pi_2 X_i + \nu_{i2},$$

with $\pi_1 = (\beta_1/(1-\gamma_1\gamma_2) \quad (\gamma_1\beta_2)/(1-\gamma_1\gamma_2))'$, $\pi_2 = ((\gamma_2\beta_1)/(1-\gamma_1\gamma_2) \quad \beta_2/(1-\gamma_1\gamma_2))'$ and $X_i = (x_{i1}' \quad x_{i2}')$.

We assume that υ_{i1} and υ_{i2} have a bivariate normal distribution with zero means, variances

[4] We leave the statistical properties of the disturbance terms unspecified for the moment.

[5] The probabilities $\Pr(y_{i1}=1, y_{i2}=1)$, $\Pr(y_{i1}=0, y_{i2}=1)$, $\Pr(y_{i1}=1, y_{i2}=0)$ and $\Pr(y_{i1}=0, y_{i2}=0)$ add to one if and only if γ_1 or γ_2 is equal to zero.

one and correlation ρ. Furthermore, the joint probabilities Pr(y_{i1}=1, y_{i2}=1), Pr(y_{i1}=0, y_{i2}=1), Pr(y_{i1}=1, y_{i2}=0) and Pr(y_{i1}=0, y_{i2}=0) are denoted as P_{11}, P_{01}, P_{10} and P_{00}. Now,

$$P_{01} = Pr(y_{i1}^{*} \leq 0, y_{i2}^{*} > 0) = \Pr(\nu_{i1} \leq -\pi_1 X_i, \nu_{i2} \leq \pi_2 X_i) = \Phi_2(-\pi_1 X_i, \pi_2 X_i, -\rho)$$

with $\Phi_2(\cdot,\cdot,\rho)$ the bivariate normal distribution function (with variances one)[6]. Estimates of π_1, π_2 and ρ can be obtained by maximizing the log-likelihood $\sum_{i=1}^{n}\sum_{s=0}^{1}\sum_{t=0}^{1} I(y_{i1}=s, y_{i2}=t)\ln P_{st}$, where $I(\cdot)$ denotes the familiar indicator function. In practice, variables can appear only once in X_i; therefore, alternative vectors π_1^* and π_2^* will be estimated. The (new) vector of explanatory variables is denoted as X_i^* which contains all available variables including an intercept (length k+1).

For variables that appear in both x_{i1} and x_{i2} (in total, k-2 variables) estimates can be calculated using the following equalities (note that estimates of γ_1, γ_2, β_1 and β_2 are denoted with a hat):

$$(\hat{\beta}_1 + \hat{\gamma}_1\hat{\beta}_2)/(1-\hat{\gamma}_1\hat{\gamma}_2) = \hat{\pi}_1^*;$$
$$(\hat{\beta}_2 + \hat{\gamma}_2\hat{\beta}_1)/(1-\hat{\gamma}_1\hat{\gamma}_2) = \hat{\pi}_2^*.$$

Hence, this is a system of 2k-2 equations (including the one for the intercept) in 2k+2 unknowns. For the remaining two variables one can obtain four equality restrictions resulting in 2k+2 equations. Estimates of γ_1, γ_2, β_1 and β_2 can now easily be obtained.[7]

We approximate the standard errors of these parameters with simulation methods.[8] We know that estimates of π_1^* and π_2^* are asymptotically normal distributed as $\hat{\pi}_1^* \sim N(\pi_1^*, I(\pi_1^*)^{-1})$ and $\hat{\pi}_2^* \sim N(\pi_2^*, I(\pi_2^*)^{-1})$ with $I(\pi_1^*)$ and $I(\pi_2^*)$ the information matrices and $N(\cdot)$ the normal density. We have the availability of estimates of $I(\pi_1^*)^{-1}$ and $I(\pi_2^*)^{-1}$, denoted with $\hat{V}(\pi_1^*)$ and $\hat{V}(\pi_2^*)$. One can now simulate $\pi_1^{*(m)} \sim N(\hat{\pi}_1^*, \hat{V}(\pi_1^*))$ and $\pi_2^{*(m)} \sim N(\hat{\pi}_2^*, \hat{V}(\pi_2^*))$ for $m = 1,\ldots,M$.[9] The relationship between $\hat{\pi}_1^*$ and estimates of $\beta_1^*, \beta_2^*, \gamma_1$ and γ_2 is given by $\hat{\pi}_1^* = f(\hat{\beta}_1^*, \hat{\beta}_2^*, \hat{\gamma}_1, \hat{\gamma}_2)$. One could now reverse this relationships such that, for example, $\hat{\beta}_1^* = f(\hat{\pi}_1^*, \hat{\beta}_2^*, \hat{\gamma}_1, \hat{\gamma}_2)$. Next, draws of β_1 can be obtained since $\hat{\beta}_1^{*(m)} = f(\hat{\pi}_1^{*(m)}, \hat{\beta}_2^*, \hat{\gamma}_1, \hat{\gamma}_2)$ for $m = 1,\ldots,M$. One can now determine $\alpha/2$ per cent and $1-\alpha/2$ per cent quantiles for $\alpha = 10$, $\alpha = 5$ and $\alpha = 1$. In each of these three cases zero might be included in the confidence region. If zero is not included in the region for $\alpha = 10$, $\alpha = 5$ and $\alpha = 1$, one can conclude that the particular variable is significant at the α per cent level.

3.3. Pre-analysis

First, we re-estimate the two probit equations from Grilo and Thurik (2005a) with the present data set. The authors estimate the following probit models independently: $\Pr(y_{i1}=1) = \Phi(x_{i1}'\beta_1 + \gamma_1 y_{i2})$ and $\Pr(y_{i2}=1) = \Phi(x_{i2}'\beta_2)$ with $x_{i1} = x_{i2}$ denoting the same set of explanatory variables and where the symbols are as before. Estimation results (with average marginal effects) are displayed in Table 2 and the left-hand panel of Table 3.

---

[6] Note the sign change of ρ into -ρ in the expressions for P_{01} and P_{00}.

[7] Note that β_1 and β_2 each have size k (including intercept) because each of these vectors excludes one variable (see below). The correlation coefficient ρ in the reduced form system is estimated as 0.524 (significant at one per cent).

[8] One could also approximate standard errors with the delta method (see Greene, 2003). Both methods lead to the same conclusions.

[9] We set the number of simulations M equal to 10,000.

Note that in the first equation the probability of being self-employed is explained. When actual entrepreneurship equals zero, this analysis also includes respondents who once had a business and those who are thinking about starting an own business. We find significant effects for latent entrepreneurship (+), gender (+), age (+, no quadratic relationship) and perception of administrative complexities (-). Furthermore we find significant effects for self-employed parents (+), perception of insufficient information (+), risk tolerance (+) and internal and external success factors (both +). The sign of the impact of perception of insufficient info is unexpected; Grilo and Thurik (2005a) report an insignificant coefficient for risk tolerance. Note that education, perception of financial barriers and the perception of an unfavourable economic climate do not have significant influences.

Concerning latent entrepreneurship significant effects are found for gender (+), age (negative up to 56 years), perception of financial support (+), perception of administrative complexities (-) and risk tolerance (+). Furthermore, we find significant effects for self-employed parents (+), perception of insufficient information (+), perception of an unfavourable economic climate (-) and internal success factors (+), where again the sign of insufficient information is counterintuitive. No significant effects for education and external success factors are found.[10]

In the right-hand panel of Table 3, the model formulation $\Pr(y_{i2} = 1) = \Phi(x_{i2}'\beta_2 + \gamma_2 y_{i1})$ is compared with $\Pr(y_{i2} = 1) = \Phi(x_{i2}'\beta_2)$.

3.4. Full model

In this section we focus on the "full model" in which actual and latent entrepreneurship depend on each other. Note that we have to exclude at least variable from x_{i1} and x_{i2} to ensure model identification. We decide to exclude perceived lack of financial support from x_{i1} as this variable does not have a significant effect in recent literature (Grilo and Irigoyen, 2006; Grilo and Thurik, 2005a). We remark that the only study we know that investigates influences of all variables considered here on latent entrepreneurship is Grilo and Thurik (2006). These authors find significant effects for all variables, except for education, and hence this study is not decisive in this context (we use another measure of education). We decide to exclude perceived lack of sufficient information in line of three arguments. First, lack of insufficient information has a counterintuitive positive sign in Grilo and Thurik (2006). Second, we do not believe that a genuine wish to set up a business can be hindered by perceived lack of sufficient information. Third, we do not find convincing results with respect to significant influences on latent entrepreneurship (either with or without actual entrepreneurship included in the equation) in Table 3 for this variable.

The estimation results of the simultaneous bivariate probit model are given in Table 4. Glancing at the results for actual entrepreneurship we see that they do not differ much with respect to those obtained in Table 2. Education now has a negative significant effect (zero is not contained in the 90% confidence region).

Turning to latent entrepreneurship we do not only see again the absence of significance of perceived administrative complexities, but also perceived lack of financial constraints does not have a significant influence in this context. Clearly, the inclusion of actual entrepreneurship in the latent equation gives rise to the conclusion that three perception variables (lack of financial support, administrative complexities and insufficient info) do not directly influence latent

[10] Also, a bivariate analogue is estimated in which the two probit equations are correlated with each other (intercepts are included in each equation). Surprisingly, the correlation coefficient is not significant at any relevant significance level. Furthermore the significant effect of latent entrepreneurship on actual entrepreneurship vanishes. Estimation results are available upon request from the author.

entrepreneurship but primarily have an effect on this variable *via* actual entrepreneurship. If actual entrepreneurship is omitted from the equation perception of administrative complexities and perception of insufficient information become significant determinants. Note that financial constraints do not seem to be important in either equation.

Hence, perception of administrative complexities and perception of insufficient info are important determinants of the probability of being self-employed with the preference for self-employment primarily influenced by actual entrepreneurship. The same holds true – surprisingly – for having self-employed parents. No direct influence of this variable on the willingness to become an entrepreneur is found; however, having at least one self-employed parent has an indirect effect on latent entrepreneurship via actual entrepreneurship.

Further results from Table 4 can be summarized as follows. Gender has a positive influence on both actual and latent entrepreneurship, as expected. The relationship of age is quadratically[11] for latent entrepreneurship while only linear for actual entrepreneurship. Education does not seem to be a very important factor in this context as it reveals no significance in the latent equation and a hardly significant one in the actual equation (at 10 per cent). Risk tolerance is found to significantly influence actual entrepreneurship; this is not found in Grilo and Thurik (2005a).

4. Constructing the multinomial framework

Four categories can be distinguished on the basis on the revealed preference for self-employment ("latent entrepreneurship") and the employment status at the moment the survey was conducted ("actual entrepreneurship"). Category 1 consists of individuals who are neither self-employed nor preferring to become self-employed. In category 2 respondents are collected who are not self-employed but prefer becoming self-employed. Self-employed people, but not preferring, are in category 3 and finally individuals both having and preferring an independent status are in category 4. The distribution of the number of observations across these categories is given in Table 5. One could state that category 2 and category 3 consist of people who are unsatisfied with their current situation, with category 2 containing "frustrated entrepreneurs" and category 3 containing "forced entrepreneurs".

In Grilo and Thurik (2005a) the concept of engagement levels is introduced to discriminate between the various stages of setting up or closing down a business. The engagement levels range from "Never thought about it" to "Thinking about it", "Gave up", "Taking steps", "Having a young business", "Having a mature business" and "No longer having a business". The distribution of the engagement levels across the four categories and across respondents that are categorized as actual entrepreneurs can be read from Table 6.

Because of the categorization we make use of a multinomial logit model to give insight in the characteristics of the individuals in the four categories and the major differences between the individuals. The multinomial logit model begins by assuming that the unordered dependent variable Y_i takes J values with

$$\Pr(Y_i = j) = \frac{\exp(Z_i'\delta_j)}{1 + \sum_{l=1}^{J-1} \exp(Z_i'\delta_l)} \text{ for } j = 1, \ldots, J-1$$

[11] We have to be cautious with the finding of this quadratic relationship as the turning point at which the effect of age becomes positive amounts to 70 years. After deleting the quadratic term we find a negative linear relationship of age with latent entrepreneurship.

and $\Pr(Y_i = J) = (1 + \sum_{l=1}^{J-1} \exp(Z_i'\delta_l))^{-1}$ with J the "base category" (parameters corresponding to this base category are zero for model identification). Note that Z_i is a $(k+1)\times 1$ vector of explanatory variables (with k the number of explanatory variables) for individual i including the scalar one and δ_j is a $(k+1)\times 1$ category specific vector of parameters for category j.

With respect to the interpretation of the multinomial logit model one can rely on (log)odds ratios. In our case the base category is the category "neither actual nor latent" (*J*=1). The log-odds ratio is now equal to $\log(\Pr(Y_i = j)/\Pr(Y_i = J)) = Z_i'\delta_j$ for each category *j*. These ratios can be interpreted because they are linear functions of the explanatory variables. The parameters in δ_j measure the change in the log-odds ratio for a one unit change of the corresponding variable in Z_i, ceteris paribus all other variables. If the specific parameter is larger than zero, an increase in the specific variable implies that individuals are more likely to choose category *j* compared to the base category, while keeping all other variables equal. Of course, one can take other base categories than *J* to investigate effects on odds ratios one is interested in.

4.1. Interpretation

Estimation results of the multinomial logit model are displayed in Table 7 while some fit statistics can be read from Table 8.[12] We give an interpretation of the estimation results for four pairs of categories. The corresponding log-odds ratios for each variable are exhibited in Table 7. These four pairs are the most interesting and most straightforward ones to investigate and interpret. The discussion of the odds of category 3 versus 1 gives rise to the analysis of determinants of necessity entrepreneurship. With the multinomial logit framework we are able to identify some significant distinguishing characteristics of these necessity entrepreneurs. There has not been much scholarly work on the analysis of differences between opportunity and necessity (nascent) entrepreneurs. [13] Studies have focused on different aspects of opportunity and necessity entrepreneurship. Wagner (2005) investigates nascent entrepreneurs. Minniti *et al.* (2005) find that more women engage in necessity entrepreneurship than men do. Evidence has been found that lower educated people are more likely to engage in necessity entrepreneurship with unemployment acting as a so-called push factor for self-employment (Evans and Leighton, 1990). Wagner (2005) concludes that fear of failure is double as high among necessity entrepreneurs than among opportunity entrepreneurs.

The comparison of category 4 versus 2 investigates the determinants of actual entrepreneurship conditional on a declared preference for self-employment. The analyses of category 2 versus 1 and category 4 versus 3 both involve the determinants of latent

[12] For each variable the null hypothesis is tested whether the coefficients of this variable across the three categories equal zero (Wald test). Abstaining from country dummies, variables for which the null hypothesis is not rejected at one per cent are age squared, education and education squared. We also conduct Wald tests to test whether combining certain categories can be justified. Results show that no categories can be combined (combination 3 and 4 gives the least convincing result). For these two categories, again not considering country dummies, only for gender, risk tolerance, perception of economic climate and external success factors the null hypothesis of equal coefficients is rejected at 1 per cent.
It is not straightforward to show a measure of explanatory power for each category in the multinomial logit model. One could compute the average predicted probability for each category. In this way we can examine how well the model is able to identify the different categories. Predictions can be based on 1) the observations representing the specific category, or 2) on all 7,777 observations. Either result show that category 1 is identified best. Results are available upon request from the authors.
With different base categories we find evidence that the IIA ("independence of irrelevant alternatives") assumption has not been violated.

[13] Reynolds *et al.* (2002) explicitly distinguish between "opportunity-based" and "necessity-based" entrepreneurship. Opportunity-based entrepreneurs are those who choose to start an own business by taking advantage of an entrepreneurial opportunity. Necessity-based entrepreneurship involves those who start their own business because other employment options are either absent or unsatisfactory.

entrepreneurship (with a different actual status).

Odds of category 3 relative to 1 (not latent, actual/not actual)

In this paragraph we discuss the determinants of the odds of being in category 3 (not latent, actual) relative to category 1 (not latent, not actual). In both cases individuals do not have a preference for self-employment but they differ in their employment status at the moment of the survey. Because of the fact that necessity entrepreneurs are likely to be in category 3, this analysis essentially comes down at investigating the determinants of becoming an entrepreneur because of necessity.

Gender has a significant effect on the odds of 3 versus 1. In fact, being a man increases the likelihood of being in any category relative to 1 (in general, females are most likely to be present in category 1 which means that gender is a significant factor in driving away from the most passive entrepreneurial situation possible). The estimated parameter with base category 1 corresponding to gender – that is, 0.399 – measures the expected change in the log-odds ratio $\log(\Pr(Y_i = 3)/Pr(Y_i = 1))$ if a specific individual corresponds to a man, while keeping all other variables equal. Hence, men are exp(0.399) times more likely to be in category 3 than in 1.

Because of the insignificance of age a clear relationship between age and becoming self-employed for necessity entrepreneurs does not exist. If one does not want to be self-employed, clearly it does not hold that "entrepreneurship comes with age", but the decision to become an entrepreneur mainly comes down to an instantaneous decision due to a sudden change in the employment status.

Note that education has a negative sign. Clearly, the lower the education level case, the fewer other job opportunities are available and hence, the more likely it is one becomes an entrepreneur because of necessity.[14]

Apparently, having self-employed parents is one of the major drivers in being self-employed given that respondents have no preference for self-employment. It could for example be that individuals that do not want to be(come) active in the entrepreneurial world are forced to be so. This dynamic effect implies that having at least one self-employed parent makes respondents move from category 1 to category 3.[15]

Another important result is the fact that acknowledging administrative complexities to be a barrier in becoming an entrepreneur has a significant negative influence on being in 3 relative to 1. Clearly, given that one does not prefer to be(come) self-employed, the experience of these administrative complexities in affecting one's decision becoming self-employed, leads to the fact that people abstain from being self-employed. Note that this result can refer to the personal situation or to the general attitude towards entrepreneurship.

Note that the perception of a lack of financial support clearly is no barrier for non-latents to become self-employed. Clearly, respondents from both categories cannot be distinguished on basis of this perception variable On the other hand, necessity entrepreneurs experienced problems with obtaining sufficient information (significant influence for this variable), but still became an entrepreneur and hence, this variable cannot be seen as an obstacle.

It is no surprise that external success factors play a role. Clearly, the difference of individuals belonging to either of the two categories can be explained by the fact that one thinks becoming entrepreneur is due to the conviction that external factors play a role in running a

[14] Note that this variable does not play an important role in the entire analysis because of the lack of significance across the board.

[15] All odds relative to the first category are significant with respect to this variable, so clearly self-employed parents significantly distinguish the first category from the other three. It is, therefore, unlikely that an individual with self-employed parents is neither willing to become entrepreneur nor being active in the entrepreneurial world.

business successfully. (One is apparently conscious of the fact that one is 'forced' to become self-employed and one is, hence, not convinced that results can be influenced by the individual.)

The odds of any European country is larger than for the US. Any odds relating category 3 to another category is higher for Europe than for the US. This means that in European countries the prevalence of necessity entrepreneurs is larger, compared to the US.

Odds of category 4 relative to 2: (latent, actual/not actual)

In the present paragraph we determine what factors make individuals statistically different between category 4 and category 2. The difference between the two categories again does not lie in the willingness to be(come) entrepreneur, but in the actual status.

As in the previous paragraph, gender, self-employed parents (odds ratio of this variable is lower now) and perception of administrative complexities are significantly present.

Now, age has a significant effect, while the effect was absent in the previous situation. Clearly, older people are more likely to become an entrepreneur, given that one prefers becoming self-employed.[16] When we do not include education squared in our analysis education has a significant negative effect.

While in the previous case the difficulty to obtain sufficient info was significantly present, it loses its significance in this case. Probably respondents who are already willing to become self-employed do not find it difficult (or have no problems with the difficulties) to gain appropriate information or do not need the desired information.

The odds of being in category 4 rather than in 2 is not significantly affected by external factors, but (as expected) by internal factors. Clearly when one is already willing to become an entrepreneur one is apparently convinced by his or her own capacities and therefore believes that internal factors (has) play(ed) a significant role in reaching the self-employed status.

In sum, only relevant variables distinguishing frustrated respondents in category 2 from becoming self-employed are gender, age (quadratic function), self-employed parents, perception of administrative complexities and internal success factors. The very obstacle for respondents in category 2 from becoming self-employed seems to be the perception of administrative complexities (or they think they will be a problem, perhaps because of the information from their self-employed parents).

Odds of category 4 relative to 3: (actual, latent/not latent)

Within these two categories, respondents are self-employed, but their willingness to have an independent status differs between them. We can now see whether actual status is independent of the latent status in the sense that variables may differ in their influences conditional on the declared preference for self-employment.

Gender has a strong positive significant effect on the odds of being in the fourth category versus the third one. So, women are more likely to be in this third category (not latent, actual) than the fourth one (latent, actual). Men apparently have a higher preference (are more "satisfied") for their independent status than women, perhaps because the fact that women who became self-employed because of necessity.

Age does not significantly distinguish category 4 from 3. Individuals in the third and fourth category are indistinguishable with respect to self-employed parents, while all other odds are significantly influenced by this variable. Furthermore, variables that explain the differences between these two categories are the perception of the economic climate, risk tolerance and

[16] The point at which the effect of age on the log-odds ratio becomes negative is not contained within the data range. After deleting the quadratic term, we find a positive linear relationship.

external success factors.

Odds of category 2 relative to 1: (not actual, latent/not latent)

Age is negatively influencing the willingness to become an entrepreneur, given that one is not active in the entrepreneurial world (with a U-shaped relationship[17]).

Note that the perception of complexity of administrative procedures does not have discriminating power in this comparison, while self-employed parents are of significant influence. Apparently, people are not yet aware of the possible administrative complexities, while they think that financial constraints play an important role in setting up a business. It could also be that there are respondents who experienced lack of financial support as a barrier and once were entrepreneurs, but are no longer active in this world. Same arguments can be given for lack of sufficient information.

5. Evaluation of methods used in preceding chapters

In this section we intend to relate the models of the two preceding chapters, i.e., the bivariate probit model and the multinomial logit model. Also, we investigate whether (and if yes, when) the multinomial logit model has an advantage over the bivariate probit model, or vice versa.

Recall that in the multinomial logit model the probabilities are given by $\Pr(Y_i = j) = \frac{\exp(Z_i'\delta_j)}{1+\sum_{l=1}^{J-1}\exp(Z_i'\delta_l)}$ for $j = 1,\ldots,J-1$ and $\Pr(Y_i = J) = (1+\sum_{l=1}^{J-1}\exp(Z_i'\delta_l))^{-1}$ with J the base category. Note that the odds-ratio of category j versus k is equal to $\Pr(Y_i = j)/\Pr(Y_i = k) = \exp(Z_i'(\delta_j - \delta_k))$.

Next, we rewrite the probabilities as $\Pr(Y_i = 1) = \Pr_{ml}(\overline{A}_i \cap \overline{L}_i)$, $\Pr(Y_i = 2) = \Pr_{ml}(\overline{A}_i \cap L_i)$, $\Pr(Y_i = 3) = \Pr_{ml}(A_i \cap \overline{L}_i)$ and $\Pr(Y_i = 4) = \Pr_{ml}(A_i \cap L_i)$. These expressions are thus formulated as statistical events for each individual i. Note that A_i denotes actual entrepreneurship, that L_i denotes latent entrepreneurship and that $\overline{A}_i$ and $\overline{L}_i$ are the complements of these events, respectively.

It follows that we now have the availability of the joint distribution function of latent and actual entrepreneurship. One can easily derive unconditional (marginal) and conditional probabilities from this joint distribution. It is straightforward to show what the expressions for $\Pr_{ml}(A_i)$ and $\Pr_{ml}(L_i)$ are, but also the conditional probabilities $\Pr_{ml}(A_i \mid L_i)$ and $\Pr_{ml}(L_i \mid A_i)$ can be derived easily. We now derive both conditional probabilities inducing from the multinomial logit model:

$$E_{ml}(A_i \mid L_i) = \Pr_{ml}(A_i \mid L_i) = \frac{\Pr_{ml}(A_i \cap L_i)}{\Pr_{ml}(L_i)} = \frac{\Pr_{ml}(A_i \cap L_i)}{\Pr_{ml}(\overline{A}_i \cap L_i) + \Pr_{ml}(A_i \cap L_i)} = \frac{\exp(Z_i'\delta_4)}{\exp(Z_i'\delta_2) + \exp(Z_i'\delta_4)}.$$

After differentiating with respect to variable x_{ik} we obtain the following expression:

$$\frac{\partial \Pr_{ml}(A_i \mid L_i)}{\partial x_{ik}} = \frac{\exp(Z_i'(\delta_2 + \delta_4))(\delta_{k4} - \delta_{k2})}{(\exp(Z_i'\delta_2) + \exp(Z_i'\delta_4))^2}$$

[17] The turning point comes down to 59 years.

with β_{kl} the coefficient belonging to x_{ik}. A similar expression can be obtained for $E_{ml}(L_i \mid A_i) = \Pr_{ml}(L_i \mid A_i)$. Note that for a dummy variable x_{ik} the probabilities $\Pr_{ml}(A_i \mid L_i, x_{ki} = 1) - \Pr_{ml}(A_i \mid L_i, x_{ki} = 0)$ and $\Pr_{ml}(L_i \mid A_i, x_{ki} = 1) - \Pr_{ml}(L_i \mid A_i, x_{ki} = 0)$ have to be calculated. All probabilities should be averaged over relevant observations.

The expressions in the bivariate probit model read as $\Pr_{se}(A_i) = F(x_{i1}'\beta_1 + L_i^*)$ and $\Pr_{se}(L_i) = F(x_{i2}'\beta_2 + A_i^*)$ with the asterisk denoting the latent counterparts of A_i and L_i, respectively. These expressions also imply joint probabilities $\Pr_{se}(\overline{A}_i \cap \overline{L}_i), \Pr_{se}(\overline{A}_i \cap L_i), \Pr_{se}(A_i \cap \overline{L}_i)$ and $\Pr_{se}(A_i \cap L_i)$. Conditional probabilities as in the multinomial logit model (and their derivatives with respect to explanatory variables) could of course also be derived, for example

$$E_{se}(A_i \mid L_i) = \Pr_{se}(A_i \mid L_i) = \frac{\Pr_{se}(A_i \cap L_i)}{\Pr_{se}(L_i)} = \frac{\Phi_2\left(\frac{1}{1-\gamma_1\gamma_2}x_{i1}'\beta_1 + \frac{\gamma_1}{1-\gamma_1\gamma_2}x_{i2}'\beta_2, \frac{\gamma_2}{1-\gamma_1\gamma_2}x_{i1}'\beta_1 + \frac{1}{1-\gamma_1\gamma_2}x_{i2}'\beta_2, \rho\right)}{\Phi\left(\frac{\gamma_2}{1-\gamma_1\gamma_2}x_{i1}'\beta_1 + \frac{1}{1-\gamma_1\gamma_2}x_{i2}'\beta_2\right)}$$

$$= \frac{\Phi_2(\pi_1 X_i, \pi_2 X_i, \rho)}{\Phi(\pi_2 X_i)},$$

where the symbols are as before. In a similar way one can see that

$$E_{se}(L_i \mid A_i) = \frac{\Phi_2(\pi_1 X_i, \pi_2 X_i, \rho)}{\Phi(\pi_1 X_i)}.$$

For each variable x_{ik} we are now able to compute marginal conditional probabilities $\frac{\partial \Pr_{se}(A_i \mid L_i)}{\partial x_{ik}}$ and $\frac{\partial \Pr_{se}(L_i \mid A_i)}{\partial x_{ik}}$. With the expressions

$$\frac{\partial \Phi_2(c_1, c_2, \rho)}{\partial c_1} = \phi(c_1)\Phi\left(\frac{c_2 - \rho c_1}{\sqrt{1-\rho^2}}\right)$$

$$\frac{\partial \Phi_2(c_1, c_2, \rho)}{\partial c_2} = \phi(c_2)\Phi\left(\frac{c_1 - \rho c_2}{\sqrt{1-\rho^2}}\right)$$

one is able to calculate the desired derivatives ($\phi(\cdot)$ represents the standard normal density). Note that for a dummy variable x_{ik} we just calculate the probabilities $\Pr_{se}(A_i \mid L_i, x_{ki} = 1) - \Pr_{se}(A_i \mid L_i, x_{ki} = 0)$ and $\Pr_{se}(L_i \mid A_i, x_{ki} = 1) - \Pr_{se}(L_i \mid A_i, x_{ki} = 0)$. All these probabilities should again be averaged over observations contained in the estimation sample.

To have an idea of the similarity of both models, we first calculate predicted probabilities $N^{-1}\sum_i \hat{\Pr}_{ml}(A_i \mid L_i)$ and $N^{-1}\sum_i \hat{\Pr}_{ml}(L_i \mid A_i)$ on the one hand and $N^{-1}\sum_i \hat{\Pr}_{se}(A_i \mid L_i)$ and $N^{-1}\sum_i \hat{\Pr}_{se}(L_i \mid A_i)$ on the other hand, where *N* equals the size of the estimation sample. The results are given in Table 9. In investigating which model fits the data best, one can make use of a test proposed by Vuong (1989), which is able to compare two non-nested models. Doing this

for actual entrepreneurship for example, one has to calculate values $h_i = \ln(\frac{\hat{\Pr}_{ml}(A_i \mid L_i)}{\hat{\Pr}_{se}(A_i \mid L_i)})$ where the statistic itself equals $V = \frac{\sqrt{N}\bar{h}}{s_h}$ with $\bar{h}$ and s_h denoting the sample mean and sample standard error of h_i. It turns out that the bivariate probit model is preferred in both cases.

The partial derivatives of all variables are exhibited in Table 10 for the multinomial logit model and in Table 11 for the bivariate probit model.

An advantage of the multinomial model is that it makes it possible to perform group-wise analyses. Also, it is easier to calculate conditional probabilities and marginal probabilities from joint probabilities compared to the bivariate probit model. Moreover, it is possible to take into account the importance of latent entrepreneurship but without explicitly including latent entrepreneurship in the set of explanatory variables. The multinomial logit model, on the other hand, has as drawback that interpretation should always be done relative to a certain base category making interpretation probably somewhat more difficult than in case of the bivariate probit model. And, the analysis results in an abundance of tables with log-odds ratios one has to interpret. Above all, the bivariate probit model permits a causal interpretation that is not possible in the multinomial logit model due to the very nature of the model. The bivariate probit model also serves as a solution when decisions on variables are made in different points in time. In sum, each model has its own advantages and disadvantages. Both models should be used for distinct purposes. In making group-specific comparisons the multinomial logit model seems to be more appropriate while the bivariate probit model demonstrates to be more suitable when the the emphasis lies on causal effects and different points in time.

6. Conclusion

The starting point of this study is a bivariate probit model, which differs from earlier methodologies on investigating the interrelation between latent and actual entrepreneurship. By making the declared preference for self-employment (as an approximation of latent entrepreneurship) dependent on actual entrepreneurship we intend to be as complete as possible in investigating the determinants of this declared preference. We find that perceived lack of financial support, perceived administrative complexities and perceived insufficient information do not directly influence latent entrepreneurship but primarily have an effect on this variable via actual entrepreneurship. The same phenomenon can be observed for having at least one self-employed parent. A result that is in line with earlier research is that the level of education does not significantly influence the willingness to be(come) self-employed.

The observation that in future research other social and/or cultural variables should be considered as determinants or controlling factors of latent entrepreneurship is inspired by the fact that having self-employed parents does not directly influence latent entrepreneurship according to the results in this chapter.

In a next approach, we categorized all respondents into four categories based on their revealed preference for self-employment and their self-employment status at the moment of the survey. Results show that variables that have significant influences on self-employment are gender (positive influence), education (negative influence), self-employed parents, administrative complexities (negative) and locus of control (positive). These relationships do not depend on the declared preference for self-employment. While the perception of administrative complexities is an important factor in distinguishing actual entrepreneurs from those without an independent status, for lack of financial support and risk tolerance a clear

discriminating power is absent. Necessity entrepreneurs are more pessimistic about the environment compared to non-necessity entrepreneurs. Furthermore, the two types of entrepreneurs differ in the influence of either internal or external locus of control on the self-employment decision.

Both methodologies (bivariate probit model and multinomial logit model) provide new and extensive insights in the analysis of determinants of latent and actual entrepreneurship in Europe and the US. Concentrating on conditional probabilities, both models arrive at roughly similar marginal effects of explanatory variables. The multinomial model enables researchers to perform group-wise analyses, while the bivariate probit model makes is possible to take into account the importance of latent entrepreneurship without explicitly including latent entrepreneurship in the set of explanatory variables.

7. References

Blanchflower, D.G., Oswald, A. and A. Stutzer, 2001, 'Latent entrepreneurship across nations', *European Economic Review*, 45(4–6), 680–691.

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

Greene, W.H., 2003, *Econometric Analysis*, Fifth edition, Prentice-Hall International.

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

Grilo, I. and A.R. Thurik, 2006, 'Entrepreneurship in the old and the new Europe'. In: Santarelli, E. (ed.), *Entrepreneurship, Growth, and Innovation: the Dynamics of Firms and Industries (International Studies in Entrepreneurship),* Springer Verlag.

Grilo, I. and A.R. Thurik, 2008, 'Determinants of entrepreneurial engagement levels in Europe and the US', *Erasmus Research Institute for Management Report*, forthcoming.

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

Maddala, G.S., 1983, *Limited-dependent and qualitative variables in econometrics*, Cambridge: Cambridge University Press.

Masuda, T., 2006, 'The Determinants of Latent Entrepreneurship in Japan', *Small Business Economics*, 26(3), 227-240.

Minniti, M., Arenius, P. and N. Langowitz, 2005, *Global Entrepreneurship Monitor: 2004 Report on Women and Entrepreneurship*, Centre for Women's Leadership at Babson College, London Business School.

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

Reynolds, P.D., W.D. Bygrave, E. Autio, L.W. Cox, and M. Hay, 2002, *Global Entrepreneurship Monitor: 2002 Executive Report*, Babson College; London Business School and Kauffman Foundation.

Vuong, Q.H., 1989, 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica*, 57, 307–333.

Wagner, J., 2005, 'Nascent necessity and opportunity entrepreneurs in Germany: Evidence from the Regional Entrepreneurship Monitor (REM)', *Working Paper Series in Economics No. 10*, University of Lüneburg.

Table 1: Actual and latent entrepreneurship across countries (percentages of number of respondents in each country).

	observ.	actual	latent	act./lat.
Belgium	726	19.3	37.6	0.51
Denmark	294	14.4	40.5	0.35
Germany	791	18.8	45.9	0.41
Greece	916	42.3	53.3	0.79
Spain	623	17.7	62.4	0.28
France	772	9.5	44.6	0.21
Ireland	345	22.6	60.9	0.37
Italy	882	20.8	55.3	0.38
Luxembourg	383	9.8	51.2	0.19
Netherlands	677	20.3	35.6	0.57
Austria	265	21.3	42.6	0.50
Portugal	693	18.9	65.7	0.29
Finland	344	24.7	29.7	0.83
Sweden	313	14.4	38.3	0.38
United Kingdom	643	19.1	46.3	0.41
Cyprus	356	26.1	64.6	0.40
Czech Republic	648	20.7	36.6	0.56
Estonia	239	16.1	52.7	0.31
Hungary	623	20.5	50.2	0.41
Latvia	366	8.8	47.0	0.19
Lithuania	347	12.4	59.9	0.21
Malta	310	14.4	47.7	0.30
Poland	749	28.2	55.0	0.51
Slovakia	323	22.2	35.3	0.63
Slovenia	349	10.7	39.0	0.28
United States	644	19.0	67.1	0.28
Total	13,621	19.8	49.8	0.40

Table 2: Determinants of actual entrepreneurship: a probit estimation.

	coeff.	effect
Latent	0.941 ***	0.218
Gender	0.253 ***	0.057
Age	0.029 ***	0.007
(Age/100) squared	-0.897	-0.204
Education	-0.018	-0.004
(Education/100) squared	1.375	0.312
Self-employed parents	0.485 ***	0.120
Lack of financial support	-0.010	-0.002
Administrative complex.	-0.170 ***	-0.040
Insufficient info	0.094 **	0.021
Risk tolerance	0.079 **	0.018
Economic climate	0.019	0.004
Internal success factors	0.121 ***	0.028
External success factors	0.123 **	0.029
Belgium	0.313 ***	0.077
Czech Republic	0.594 ***	0.156
Denmark	0.114	0.027
Germany	0.308 ***	0.076
Estonia	0.279 *	0.069
Greece	0.827 ***	0.231
Spain	0.098	0.023
France	-0.154	-0.033
Ireland	0.148	0.035
Italy	0.227 **	0.055
Cyprus	0.351 ***	0.088
Latvia	0.019	0.004
Lithuania	0.063	0.015
Luxembourg	-0.284 *	-0.058
Hungary	0.531 ***	0.138
Malta	0.071	0.017
Netherlands	0.397 ***	0.100
Austria	0.304 **	0.076
Poland	0.522 ***	0.137
Portugal	0.055	0.013
Slovenia	-0.005	-0.001
Slovakia	0.597 ***	0.159
Finland	0.581 ***	0.155
Sweden	0.109	0.026
United Kingdom	0.245 **	0.059
Number of observations	7,777	
Log-likelihood	-3,157.812	
LR statistic	1,424.747 (χ^2, 38 df.)	
Akaike inform. crit.	0.822	
Bayesian inform. crit.	0.490	
McFadden R^2	0.184	

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

Table 3: Determinants of latent entrepreneurship: with and without the inclusion of actual entrepreneurship.

	without actual		with actual	
	coeff.	effect	coeff.	effect
Actual entrepreneurship			1.025 ***	0.355
Gender	0.373 ***	0.140	0.309 ***	0.108
Age	-0.022 ***	-0.008	-0.024 ***	-0.008
(Age/100) squared	2.303 ***	0.854	1.927 **	0.667
Education	0.001	0.000	0.006	0.002
(Education/100) squared	-0.072	-0.027	-0.543	-0.188
Self-employed parents	0.269 ***	0.100	0.136 ***	0.047
Lack of financial support	0.110 ***	0.041	0.110 ***	0.038
Administrative complex.	-0.103 ***	-0.038	-0.056	-0.019
Insufficient info	0.061 *	0.023	0.038	0.013
Risk tolerance	0.261 ***	0.098	0.239 ***	0.083
Economic climate	-0.108 ***	-0.040	-0.112 ***	-0.039
Internal success factors	0.105 ***	0.039	0.072 *	0.025
External success factors	-0.075 *	-0.028	-0.107 **	-0.037
Belgium	-0.754 ***	-0.260	-0.834 ***	-0.264
Czech Republic	-0.550 ***	-0.196	-0.697 ***	-0.226
Denmark	-0.644 ***	-0.225	-0.682 ***	-0.221
Germany	-0.415 ***	-0.150	-0.494 ***	-0.165
Estonia	-0.050 **	-0.019	-0.123	-0.042
Greece	-0.340 ***	-0.123	-0.593 ***	-0.193
Spain	-0.212	-0.078	-0.254 **	-0.087
France	-0.661 ***	-0.231	-0.642 ***	-0.210
Ireland	-0.237 **	-0.087	-0.283 **	-0.096
Italy	-0.403 ***	-0.145	-0.461 ***	-0.154
Cyprus	-0.172	-0.063	-0.260 **	-0.088
Latvia	-0.307 ***	-0.112	-0.329 ***	-0.112
Lithuania	0.070	0.026	0.034	0.012
Luxembourg	-0.357 ***	-0.130	-0.326 ***	-0.110
Hungary	-0.245 ***	-0.090	-0.376 ***	-0.127
Malta	-0.402 ***	-0.145	-0.433 ***	-0.145
Netherlands	-0.768 ***	-0.264	-0.870 ***	-0.274
Austria	-0.432 ***	-0.155	-0.516 ***	-0.171
Poland	-0.096 ***	-0.035	-0.245 **	-0.084
Portugal	-0.044	-0.016	-0.074	-0.026
Slovenia	-0.594 ***	-0.209	-0.611 ***	-0.200
Slovakia	-0.469 ***	-0.168	-0.617 ***	-0.201
Finland	-0.865 ***	-0.289	-1.026 ***	-0.310
Sweden	-0.649 ***	-0.226	-0.672 ***	-0.218
United Kingdom	-0.503 ***	-0.180	-0.561 ***	-0.185
Number of observations	7,777		7,777	
Log-likelihood	-5,040.07		4,724.01	
LR statistic	698.79 (χ^2, 38 df.)		1,330.93 (χ^2, 39 df.)	
Akaike inform. crit.	1.306		1.225	
Bayesian inform. crit.	1.341		0.720	
McFadden R^2	0.062		0.123	

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

Table 4: Estimation results simultaneous probit model (7,777 observations).

	Actual		Latent	
Latent	0.257	**		
Actual			0.564	**
Gender	0.256	***	0.175	***
Age	0.025	***	-0.032	***
(Age/100) squared	-0.656		2.283	***
Education	-0.017	*	0.011	
(Education/100) squared	1.352		-0.926	
Self-employed parents	0.460	***	-0.033	
Lack of financial support			0.091	
Administrative complex.	-0.163	***	0.007	
Insufficient info	0.090	***		
Risk tolerance	0.091	***	0.172	***
Economic climate	0.010		-0.098	***
Internal success factors	0.119	***	0.020	
External success factors	0.108	***	-0.128	***
Belgium	0.250	***	-0.795	***
Czech Republic	0.520	***	-0.775	***
Denmark	0.080		-0.606	***
Germany	0.274	***	-0.518	***
Estonia	0.271	***	-0.210	*
Greece	0.753	***	-0.727	***
Spain	0.099		-0.253	***
France	-0.160		-0.486	***
Ireland	0.140		-0.291	***
Italy	0.204	***	-0.471	***
Cyprus	0.332	***	-0.340	***
Latvia	0.025		-0.292	***
Lithuania	0.081		0.001	
Luxembourg	-0.259	**	-0.167	*
Hungary	0.489	***	-0.493	***
Malta	0.063		-0.391	***
Netherlands	0.324	***	-0.851	***
Austria	0.269	***	-0.526	***
Poland	0.496	***	-0.377	***
Portugal	0.070		-0.092	***
Slovenia	-0.022		-0.507	***
Slovakia	0.530	***	-0.708	***
Finland	0.482	***	-1.025	***
Sweden	0.072		-0.600	***
United Kingdom	0.210	**	-0.555	***

***: zero not contained in 99% confidence region; **: 95%; *: 90%.

Table 5: Latent and actual entrepreneurship: number of observations.

		latent		total
		0	1	
actual	0	3,634 (cat. 1)	2,603 (cat. 2)	6,237
	1	321 (cat. 3)	1,219 (cat. 4)	1,540
total		3,955	3,822	7,777

Table 6: Distribution of four categories over engagement levels (number of observations and percentages).

Engagement level	Category 1		Category 2		Category 3		Category 4	
	abs.	%	abs.	%	abs.	%	abs.	%
"Never thought about it"	2,485	68	881	34	68	21	154	13
"Thinking about it"	510	14	1,016	39	18	6	70	6
"Gave up"	335	9	291	11	11	3	21	2
"Taking steps"	28	1	177	7	11	3	67	6
"Young business"	26	1	50	2	71	22	269	22
"Mature business"	39	1	55	2	123	38	604	50
"No longer"	211	6	133	5	19	6	34	3
All categories	3,634		2,603		321		1,219	

Engagement level	Actual=0		Actual=1	
	abs.	%	abs.	%
"Never thought about it"	3,366	54	222	14
"Thinking about it"	1,526	24	88	6
"Gave up"	626	10	32	2
"Taking steps"	205	3	78	5
"Young business"	76	1	340	22
"Mature business"	94	2	727	47
"No longer"	344	6	53	3
All categories	6,237		1,540	

Table 7: Estimation results multinomial logit (log-odds ratios).

	3 versus 1	4 versus 2	4 versus 3	2 versus 1
Latent	N	Y	Y/N	Y/N
Actual	Y/N	Y/N	Y	N
Gender	0.399 ***	0.478 ***	0.575 ***	0.496 ***
Age	0.003	0.071 ***	0.020	-0.048 ***
(Age/100) squared	2.987	-3.469 *	-2.418	4.038 **
Education	-0.078 *	-0.004	0.062	-0.013
(Education/100) squared	6.810	-1.737	-5.897	2.649
Self-employed parents	1.071 ***	0.749 ***	-0.021	0.301 ***
Lack financial support	0.264	-0.126	-0.154	0.237 ***
Administrative complex.	-0.282 *	-0.316 ***	-0.124	-0.090
Insufficient info	0.362 ***	0.073	-0.181	0.108 *
Risk tolerance	0.140	0.127	0.380 ***	0.393 ***
Economic climate	0.271 *	-0.019	-0.443 ***	-0.153 **
Internal success factors	0.153	0.247 ***	0.199	0.106
External success factors	0.398 **	0.120	-0.401 **	-0.123
Belgium	1.476 *	0.728 ***	-2.162 ***	-1.414 ***
Czech Republic	1.972 **	1.192 ***	-1.967 **	-1.187 ***
Denmark	1.147	0.250	-1.976 **	-1.079 ***
Germany	1.583 **	0.543 ***	-1.815 **	-0.775 ***
Estonia	2.137 **	0.260	-1.956 **	-0.079
Greece	3.150 ***	1.074 ***	-2.666 ***	-0.591 ***
Spain	1.786 **	-0.021	-2.095 ***	-0.288 *
France	1.330 *	-0.618 **	-2.874 ***	-0.926 ***
Ireland	1.372	0.235	-1.561 *	-0.423 **
Italy	2.145 ***	0.050	-2.651 ***	-0.556 ***
Cyprus	1.743 **	0.510 **	-1.602 *	-0.370 *
Latvia	1.760 **	-0.471	-2.633 ***	-0.402 **
Lithuania	1.537	-0.042	-1.442	0.138
Luxembourg	0.719	-0.534 *	-1.725 *	-0.472 **
Hungary	1.887 **	0.980 ***	-1.518 *	-0.611 ***
Malta	1.720 **	-0.114	-2.417 ***	-0.584 ***
Netherlands	1.820 **	0.769 ***	-2.486 ***	-1.434 ***
Austria	0.479	0.837 ***	-0.619	-0.976 ***
Poland	2.780 ***	0.579 ***	-2.318 ***	-0.118
Portugal	1.666 **	-0.032	-1.717 **	-0.018
Slovenia	1.190	-0.059	-2.182 **	-0.934 ***
Slovakia	1.576 *	1.319 ***	-1.387	-1.129 ***
Finland	1.855 **	1.474 ***	-2.344 ***	-1.963 ***
Sweden	1.556 *	0.069	-2.514 ***	-1.026 ***
United Kingdom	1.333 *	0.492 **	-1.770 **	-0.929 ***
Constant	-4.865 ***	-3.788 ***	2.218 *	1.141 ***

Table 8: Diagnostic measures for multinomial logit model.

Number of observations	7,777
Log-likelihood	-8,114.538
LR statistic (χ^2_{114})	1,563.004
McFadden R^2	0.088
Nagelkerke R^2	0.203
Akaike inform. crit.	2.117
Bayesian inform. crit.	3.350

Table 9: Average predictions of both models (based on 8,540 observations); standard errors between parentheses.

$N^{-1}\sum_i \hat{\Pr}_{ml}(A_i \mid L_i)$	0.314 (0.163)	$N^{-1}\sum_i \hat{\Pr}_{se}(A_i \mid L_i)$	0.310 (0.149)
$N^{-1}\sum_i \hat{\Pr}_{ml}(L_i \mid A_i)$	0.788 (0.117)	$N^{-1}\sum_i \hat{\Pr}_{se}(L_i \mid A_i)$	0.782 (0.108)

Table 10: Partial derivatives multinomial logit model (8,540 observations).

	A | *L*		*L* | *A*	
	coeff.	std.err	coeff.	std.err
Gender	0.086	0.024	0.076	0.028
Age	-0.507	0.241	-0.236	0.117
(Age/100) squared				
Education	-0.048	0.020	-0.396	0.167
(Education/100) squared				
Self-employed parents	0.140	0.033	0.000	0.000
Lack of financial support	-0.032	0.009	0.001	0.000
Administrative complex.	-0.064	0.017	-0.022	0.009
Insufficient info	0.010	0.003	-0.031	0.012
Risk tolerance	0.016	0.005	0.063	0.023
Economic climate	-0.004	0.001	-0.067	0.025
Internal success factors	0.043	0.012	0.041	0.016
External success factors	0.019	0.005	-0.053	0.019
Belgium	0.102	0.025	-0.095	0.032
Czech Republic	0.186	0.039	-0.064	0.022
Denmark	0.054	0.014	-0.037	0.013
Germany	-0.001	0.000	-0.059	0.021
Estonia	0.005	0.001	-0.058	0.020
Greece	0.170	0.036	-0.211	0.056
Spain	-0.048	0.015	-0.097	0.032
France	-0.136	0.048	-0.241	0.060
Ireland	0.015	0.004	-0.018	0.007
Italy	-0.033	0.010	-0.209	0.055
Cyprus	0.057	0.015	-0.009	0.003
Latvia	-0.125	0.044	-0.207	0.055
Lithuania	-0.055	0.018	0.020	0.008
Luxembourg	-0.134	0.048	-0.028	0.010
Hungary	0.144	0.032	0.014	0.006
Malta	-0.065	0.021	-0.157	0.046
Netherlands	0.101	0.024	-0.164	0.049
Austria	0.118	0.028	0.115	0.052
Poland	0.069	0.018	-0.136	0.041
Portugal	-0.048	0.015	-0.031	0.011
Slovenia	-0.055	0.018	-0.107	0.034
Slovakia	0.224	0.043	0.027	0.011
Finland	0.245	0.044	-0.133	0.041
Sweden	-0.033	0.010	-0.175	0.050
United Kingdom	0.047	0.013	-0.033	0.012

Table 11: Partial derivatives bivariate probit model (8,540 observations).

	$A \mid L$		$L \mid A$	
	coeff.	std.err	coeff.	std.err
Gender	0.082	0.017	0.067	0.018
Age	-0.327	0.281	0.709	0.290
(Age/100) squared				
Education	0.206	0.081	0.045	0.050
(Education/100) squared				
Self-employed parents	0.155	0.025	0.020	0.007
Lack of financial support	-0.059	0.012	0.005	0.005
Administrative complex.			-0.006	0.003
Insufficient info	0.016	0.002		
Risk tolerance	0.017	0.005	0.064	0.017
Economic climate	0.005	0.002	-0.036	0.010
Internal success factors	0.030	0.007	0.017	0.005
External success factors	0.033	0.007	-0.031	0.008
Belgium	0.049	0.013	-0.188	0.035
Czech Republic	0.133	0.023	-0.159	0.031
Denmark	0.036	0.009	-0.079	0.019
Germany	-0.024	0.008	-0.121	0.027
Estonia	0.032	0.007	0.030	0.009
Greece	0.231	0.030	-0.126	0.028
Spain	-0.023	0.005	0.003	0.001
France	-0.093	0.024	-0.086	0.021
Ireland	0.011	0.002	-0.012	0.003
Italy	0.021	0.006	-0.069	0.017
Cyprus	0.062	0.012	-0.014	0.004
Latvia	-0.053	0.013	-0.009	0.003
Lithuania	-0.040	0.011	0.079	0.026
Luxembourg	-0.135	0.038	0.022	0.007
Hungary	0.111	0.019	-0.050	0.013
Malta	-0.034	0.008	-0.044	0.011
Netherlands	0.072	0.017	-0.205	0.037
Austria	0.039	0.009	-0.082	0.019
Poland	0.121	0.020	-0.015	0.005
Portugal	-0.034	0.008	0.049	0.015
Slovenia	-0.055	0.014	-0.091	0.022
Slovakia	0.146	0.024	-0.131	0.028
Finland	0.135	0.025	-0.267	0.041
Sweden	-0.023	0.008	-0.124	0.028
United Kingdom	0.020	0.007	-0.098	0.023

The results of EIM's Research Programme on SMEs and Entrepreneurship are published in the following series: Research Reports and Publieksrapportages. The most recent publications of both series may be downloaded at: www.eim.net.

Recent Research Reports and Scales Papers

H200718	21-12-2007	Knowledge Management and Innovation: An empirical study of Dutch SMEs
H200717	21-12-2007	Entrepreneurship and innovation
H200716	21-12-2007	Employment Growth of New Firms
H200715	21-12-2007	Entrepreneurial Culture and its Effect on the Rate of Nascent Entrepreneurship
H200714	21-12-2007	Creative industries
H200713	19-11-2007	New Ventures' Export Orientation: Outcome And Source Of Knowledge Spillovers
H200712	29-10-2007	SME Choice of Direct and Indirect Export Modes: Resource Dependency and Institutional Theory Perspectives
H200711	24-10-2007	Family Orientation, Strategic Orientation and Innovation Performance in SMEs: A Test of Lagged Effects
H200710	15-10-2007	Drivers of entrepreneurial aspirations at the country level: the role of start-up motivations and social security
H200709	12-10-2007	Does Self-Employment Reduce Unemployment?
H200708	10-9-2007	Social security arrangements and early-stage entrepreneurial activity
H200707	11-5-2007	Competition and innovative intentions: A study of Dutch SMEs
H200706	eind maart	High-Growth Support Initiatives
H200705	14-2-2007	The relationship between economic development and business ownership revisited
H200704	2-2-2007	The relationship between knowledge management, innovation and firm performance: evidence from Dutch SMEs
H200703	26-1-2007	Family orientation, strategy and organizational learning as predictors of knowledge management in Dutch SMEs
H200702	3-1-2007	Ambitious Nascent Entrepreneurs and National Innovativeness
H200701	3-1-2007	Entrepreneurial diversity and economic growth
H200627	21-12-2006	Motivation Based Policies for an Entrepreneurial EU Economy
H200626	19-12-2006	Export Orientation among New Ventures and Economic Growth
H200625	18-12-2006	Institutionele voorwaarden voor zelfstandig ondernemerschap
H200624	13-12-2006	Creative Destruction and Regional Competitiveness
H200623	6-12-2006	Entrepreneurship, Dynamic Capabilities and New Firm Growth
H200622	1-12-2006	Determinants of self-employment preference and realization of women and men in Europe and the United States
H200621	1-12-2006	Is human resource management profitable for small firms?
H200620	23-11-2006	The entrepreneurial ladder and its determinants
H200619	20-11-2006	Knowledge Spillovers and Entrepreneurs' Export Orientation
H200618	20-11-2006	The effects of new firm formation on regional development over time: The case of Great Britain
H200617	11-10-2006	On the relationship between firm age and productivity growth
H200616	11-10-2006	Entrepreneurship and its determinants in a cross-country setting

H200615	2-10-2006	The Geography of New Firm Formation: Evidence from Independent Start-ups and New Subsidiaries in the Netherlands
H200614	25-9-2006	PRISMA-K: een bedrijfstakkenmodel voor de korte termijn
H200613	25-9-2006	PRISMA-M: een bedrijfstakkenmodel voor de middellange termijn
H200612	25-9-2006	PRISMA-MKB: modelmatige desaggregatie van bedrijfstakprognose naar grootteklasse
H200611	25-9-2006	PRISMA-R: modelmatige desaggregatie van bedrijfstakprognoses naar provincie
H200610	25-9-2006	Explaining engagement levels of opportunity and necessity entrepreneurs
H200609	25-9-2006	The effect of business regulations on nascent and Young business entrepreneurship
H200608	24-8-2006	High growth entrepreneurs, public policies and economic growth
H200607	18-8-2006	The decision to innovate
H200606	6-7-2006	Innovation and international involvement of Dutch SMEs
H200605	27-6-2006	Uncertainty avoidance and the rate of business ownership across 21 OECD countries, 1976-2004
H200604	22-6-2006	The Impact of New Firm Formation on Regional Development in the Netherlands
H200603	21-6-2006	An Ambition to Grow
H200602	21-6-2006	Exploring the informal capital market in the Netherlands: characteristics, mismatches and causes
H200601	22-5-2006	SMEs as job engine of the Dutch private economy
N200520	7-3-2006	High Performance Work Systems, Performance and Innovativeness in Small Firms
N200519	1-2-2006	Entrepreneurial Culture as Determinant of Nascent Entrepreneurship
N200518	26-1-2006	Social security arrangements and early-stage entrepreneurial activity; an empirical analysis
N200517	23-1-2006	Determinants of Growth of Start-ups in the Netherlands
N200516	23-1-2006	Entrepreneurship in the old en new Europe
N200515	23-1-2006	Entrepreneurial engagement levels in the European Union
N200514	23-1-2006	Latent and actual entrepreneurship in Europe and the US: some recent developments
N200513	20-1-2006	Determinants of self-employment preference and realisation of women and men in Europe and the United States
N200512	20-1-2006	PRISMA-K: een bedrijfstakkenmodel voor de korte termijn
N200511	19-1-2006	Strategic Decision-Making in Small Firms: Towards a Taxonomy of Entrepreneurial Decision-Makers
N200510	11-1-2006	Explaining female and male entrepreneurship at the country level
N200509	11-1-2006	The link between family orientation, strategy and innovation in Dutch SMEs: a longitudinal study
N200508	11-1-2006	From nascent to actual entrepreneurship: the effect of entry barriers
N200507	11-1-2006	Do entry barriers, perceived by SMEs, affect real entry? Some evidence from the Netherlands
H200503	6-12-2005	The Impact of New Firm Formation on Regional Development in the Netherlands