The Social Dynamics of Entrepreneurship in a Transition Economy

Michael Bräuninger* Dragos Radu†‡

February 1, 2006

Abstract

This paper studies both theoretically and empirically the determinants of entrepreneurial behaviour in a changing social environment. The theoretical setting is based on a multiple-equilibria model with endogenous costs of failure (stigma). Precisely because the case of a transition economy is a quasi-experimental setting pertinent to analyse entrepreneurial activities, the model is then estimated with a large data set on Romania: the National Demographic Survey (2003). The constructed data set allows the identification endogenous (regressed individual outcome on peer-group behaviour) social effects on entrepreneurial decisions. A maximum-likelihood probit model with sample selection is estimated to account for the self-selective nature of the decision timing of nascent-entrepreneurs. The estimated significant social effects support the predictions of the theoretical model. The results imply the presence of negative income premia and of significant non-pecuniary benefits from entrepreneurial activity (prestige).

JEL: J24, R12, M13
Keywords: entrepreneurship, social interaction, transition economy

*Economic Trends, Hamburg Institute of International Economics, Email: Michael.Braeuninger@hwwa.de
†Migration Research Group, Hamburg Institute of International Economics, Email: Dragos.Radu@hwwa.de (corresponding author)
‡Support in the context of the Marie Curie Excellence Grant: "Expanding the Knowledge Base of European Labour Migration Policies" is gratefully acknowledged.
1 Introduction

One important aspect the recent economic literature on the determinants of entrepreneurial activities hardly accounts for is the pronounced and persistent variation in entrepreneurship within and between otherwise similar countries or regions. Differentials in the prevalence rate of entrepreneurial activities are usually traced back to either differences between individuals (like e.g. the level of education or occupational skills, the family structures, risk aversion) or between formal institutions (labour market regulations, differences in tax incentives or in administrative costs). But there are few plausible reasons for why the average abilities of individuals should differ between countries or regions and why formal institutions could be different among the regions of a country with a centralised political-administrative system. Thus, if regional units are similar with regard to institutions and the characteristics of inhabitants what explains the differences in entrepreneurship?

There are only two additional possibilities left to be considered for the analysis: either differentials in local economic conditions (fundamentals) or the presence of social multipliers\(^1\) (through which aggregate relationships overstate individual elasticities).

In the context of entrepreneurial behaviour, the presence of positive social interactions or strategic complementarities refers to particular forms of externalities through which individual choices over entrepreneurship are affected by the entrepreneurial actions of a reference group. In other words, the question arises if the propensity of an individual to become an entrepreneur will vary with the prevalence of entrepreneurial behaviour in some reference group containing the individual. This will be the case if the marginal utility to one person of undertaking an action is increasing with the average amount of the action taken by his peers. Models with social multipliers seem to be particularly appropriate when analysing cross-sectional differences in entrepreneurial activity.

This paper uses a multiple-equilibria model with endogenous costs of failure (stigma) to study the determinants of entrepreneurial behaviour in a changing social environment.

The aim of the paper is to analyse to what extent the observed persistent and as yet unexplained regional differentials in entrepreneurship are due to

---

\(^1\)For an analysis of different cases of social multipliers see Edwards, Sacerdote, Scheinkman 2003.
the existence of positive social interactions.

Precisely because the case of a transition economy is a quasi-experimental setting pertinent to analyse entrepreneurial activities, the model is estimated with a large micro-data set on Romania. The constructed data set allows the identification of both endogenous (regressed individual outcome on peer-group behaviour) and contextual (regressed individual outcome on peer-group characteristics) social effects on entrepreneurial decisions. To test the robustness of endogenous social interactions two additional controls are introduced: the employment biography (Lazear’s Jack-of-all-trade) and the exposure to work experience abroad (return migration).

Methodologically there are two main challenges to be solved in the estimation. The first one is related to the identification of the social interaction parameters - Manski’s (1993) so called ”reflection problem”. This is solved due to the fact that the data are representative at regional level (allowing to define the reference groups locally) and include retrospective information on the entrepreneurial decisions of individuals (permitting thus the construction of lagged behavioural variables). The second challenge refers to the self-selective nature of the decision timing of nascent-entrepreneurs. For this reason, a maximum-likelihood probit model with sample selection is estimated to correct the selection effects.

The remainder of the paper is organised as follows. The following section offers an overview on the related literature and explains how the present paper contributes to the ongoing debate in economics. Section 3 presents the features of the modelling strategy and discusses the implications of the theoretical results. Section 4 describes the data set and provides details on the constructed variables and the empirical specification. The identification problem and the alternative chosen to overcome it are discussed in section 5. Section 6 presents the results of the estimation and their consistency with the predictions of the model. Section 7 concludes.

2 Background and related literature

There exists by now a large bulk of economic literature on the determinants of entrepreneurial activities. Recent developments in this literature try to explain the wide cross-regional differences in entrepreneurship by incorporating - besides the effects of individual characteristics and of the economic environment - the effects of the social environment, i.e. positive social interactions
or social multipliers.

There are different strands in this literature\(^2\). Usually the starting assumption is that social interactions matter because they create social norms (Akerlof 1980). Landier (2002) is one of the seminal models to explain regional differences in entrepreneurial activities across regions and sectors. In this model a "stigma of failure" enters endogenously the entrepreneurial decision making process.

The career concern literature (e.g. Holmström 1982, 1999; Bhide 2000) delivers arguments for why status and recognition matter in the occupational choices and how entrepreneurial decisions impact on reputation.

A different strand of the literature includes the variations in entrepreneurial behaviour in the context of decision making processes which take into account the effects of cultural values (e.g. Guido et al. 2003, Davidson and Wiklund 1997).

Although the theoretical modelling of these questions has progressed rapidly, there are hardly econometric studies testing the implication of the models.

One of the exceptions is Giannetti and Simonov (2005). Using a register-based longitudinal data set for Sweden they were able to show that social interactions play an important role in the decision to become an entrepreneur. Their empirical strategy is akin to that in the present paper since they also use both individual and regional level controls for the determinant of entrepreneurship and define the reference groups locally.

Apart from its contribution to the applied literature on the social effects on entrepreneurial activities, this paper also contributes to another debate in economics, namely that on entrepreneurship in the transition economies. In this context, the development of a new private sector is considered crucial to economic transition. However, until now there was only little empirical research on the nature and determinant of entrepreneurial behaviour.

While recent studies addressed the characteristics of self-employment and of business start-ups in Romania (e.g. Earle et. al 2004 and 2005) this study is the first one to suggest the existence of social effects in entrepreneurial behaviour and to provide a closer look at the regional variation of entrepreneurship in Romania.

\(^2\)For a timely and comprehensive survey compare Giannetti and Simonov 2004.
3 Theoretical model

3.1 The individual decisions on entrepreneurship

Individuals live for two periods. They maximize lifetime utility which depends on income and social prestige. All individuals differ with respect to their abilities $a_i$. In the first period they work as an employee of a firm. The income of an individual $i$ is $a_i y$. In the second period, individuals can either work as an employee or they can become an entrepreneur. If they stay as an employee their income is again $a_i y$. If they become an entrepreneur their income is $\rho a_i y$, with $\rho > 1$. However, self employment is a risky activity. With probability $\pi_i$ an individual is successful as an entrepreneur and with probability $1 - \pi_i$ the individual is not successful. Social prestige of employees and of entrepreneurs is $a_i s$. Those entrepreneurs going bankrupt have to become employed again; so their income is again $a_i y$. However, the failure in entrepreneurship leads to a loss in prestige and so prestige is below of those who never started entrepreneurship. Denote the discount factor for social prestige through failure with $\mu > 1$. Then the prestige of anfailed entrepreneur is $s_f = \mu^{-1} a_i s$.

An individual becomes self employed if expected utility of being an entrepreneur exceeds the certain utility in the case of employment. This is the case if:

$$\pi_i u(\rho a_i y, a_i s) + (1 - \pi_i) u(a_i y, \mu^{-1} a_i s) > u(a_i y, a_i s)$$

The condition is fulfilled if the probability of success exceeds a critical level:

$$\pi_i > \pi_c = \frac{u(a_i y, a_i s) - u(a_i y, \mu^{-1} a_i s)}{u(\rho a_i y, a_i s) - u(a_i y, \mu^{-1} a_i s)}$$  \hspace{1cm} (1)

The probability level is an increasing function $\mu$ and a declining function of $\rho$.

The individual probability of success depends on the ability. Income and social prestige in the different states, i.e. self employment, employed and failed self employment, will also depend on individual ability. For simplicity we assume proportionality and so - due to linear homogeneity of utility function - ability does not affect the right hand side of the condition (1). Assume a simple logarithmic utility function. In this case, expected utility of an entrepreneur in the second period can be written as:

$$\pi \gamma \log(\rho a_i y_s) + \pi (1 - \gamma) \log(a_i s) + (1 - \pi) \gamma \log(a_i y) + (1 - \pi) (1 - \gamma) \log(\mu^{-1} a_i s).$$
The second period utility of an employee is $\gamma \log(a_i y) + (1 - \gamma) \log(a_i s)$. In this case the critical probability level simplifies to:

$$\pi > \pi_c = \frac{\log(\mu)}{\log(\rho) + \log(\mu)}$$

### 3.2 The proportion of entrepreneurs

Now we assume the success probability to be an increasing function of the ability $\pi_i = (\alpha^{-1} a_i)^{1/\beta}$, where $\beta > 1$. So the critical success probability implies a critical level of ability $a_c = \alpha \pi_c^\beta$. Abilities are a random variable distributed over the range $[0, \infty]$ with a the distribution function $f(a_i)$ and a cumulated distribution function $F(a_i)$. The proportion of individuals with a success probability above the critical level is $p = 1 - F(a_c) = 1 - F(\alpha \pi_c^\beta)$.

The proportion depends on the critical probability level. For the derivatives we obtain:

$$\frac{\partial p}{\partial \pi_c} = -f(\alpha \pi^\beta)\beta \alpha \pi^{\beta-1} < 0$$

$$\frac{\partial^2 p}{\partial \pi_c^2} = -\left((\beta - 1)f(\alpha \pi^\beta)\pi^{\beta-2} - f'(\alpha \pi^\beta)\beta \alpha \pi^{2(\beta-1)}\right) \beta \alpha$$

So the proportion of entrepreneurs declines with the critical probability level. As the second order derivative shows, the function is $S$-shaped.

The critical probability level and, therefore, the self employed proportion depends on the difference in income between self employed and employed individuals and on the difference in social respect between successful and unsuccessful entrepreneurs. Both are influenced by the proportion of entrepreneurs.

To derive the wage differential, we assume a Cobb-Douglas type production function: $Y = A(pN)^{\delta}((1 - p)N)^{1-\delta}$, with $\delta > 0.5$. Both entrepreneurs and employees are paid their marginal product:

$$w_E = \frac{\partial Y}{\partial E} = \frac{\delta Y}{pN} \quad \text{and} \quad w_L = \frac{\partial Y}{\partial L} = \frac{(1 - \delta)Y}{(1 - p)N}$$

The relative wage is then given by:
\[
\rho = \frac{w_E}{w_L} = \frac{\delta Y (1 - p) N}{pN (1 - \delta) Y} = \frac{\delta (1 - p)}{(1 - \delta) p}
\]

We have \( \rho > 1 \) if \( \delta > p \). So maximum proportion of entrepreneurs is given by \( \delta \). An increase in the proportion of entrepreneurs leads to a decline in the relative wage:

\[
\frac{\partial \rho}{\partial p} = \frac{\delta}{p^2(\delta - 1)} < 0
\]

The difference in prestige between successful and failed entrepreneurs is assumed to depend on the societies experience with entrepreneurship. In particular we assume the factor \( \mu \) to depend on the proportion on those who start entrepreneurship: \( \mu = \mu(p) \) with \( \partial \mu / \partial p < 0 \). In addition we assume that the social prestige discount for failed entrepreneurs is proportional for the rewards of an successful entrepreneur. So we assume \( \mu = \eta \rho \) which gives:

\[
\pi_c = \frac{\log(\eta\rho)}{\log(\rho) + \log(\eta\rho)}
\]

Rearrange and differentiate to obtain:

\[
\frac{\partial \pi_c}{\partial p} = \frac{\log(\eta)}{(1 - p) \rho^2 (2 \log(\rho) + \log(\eta))^2}
\]

Depends on the sign of \( \log(\eta) \). If \( \eta < 1 \) we have \( \partial \pi_c / \partial p < 0 \) and with \( \eta > 1 \) we have \( \partial \pi_c / \partial p > 0 \). In the following we assume \( \eta < 1 \), which is the interesting case. This means that an increase in the proportion of skilled individuals leads to an decline in the critical probability level. The main result is: A decline in the critical probability level leads to an increase in the proportion of entrepreneurs. On the other hand, an increase in the proportion of entrepreneurs leads to a decline in the critical probability level. This is illustrated in figure 1. As a results, there will be three equilibria: One with low, one with intermediate and one with high entrepreneurial activity.

In the equilibrium with low entrepreneurial activity there are only few entrepreneurs. Close to the equilibrium there are high rewards for entrepreneurs but still people do not want to become an entrepreneur because the fear of failure is high due to high loss in social prestige in the case of failure.
Figure 1: Equilibrium and model stability

In the equilibrium with high entrepreneurial activity there are many entrepreneurs. Close to the equilibrium there are low rewards for entrepreneurs but still people do try become an entrepreneur because the fear of failure is low. Therefore, the equilibria with high and low entrepreneurial activity are locally stable. This implies that very similar societies might have very different entrepreneurial activity due to a different historical background. In the next section we will examine whether this is an observable characteristic of a society.

4 Data and empirical implementation

The paper uses data from a large demographic survey conducted 2003 by the Centre for Urban and Regional Sociology (CURS), Bucharest, on a countrywide representative sample of about 35,000 individuals. The survey collected information on a wide range of individual and household characteristics as well as information on migration experience, retrospective information on the employment biography and entrepreneurial activities.

Two main features made this survey particularly appropriate for the

\footnote{Access to the data sets have been obtained in the context of the Marie Curie Excellence Grant: Expanding the Knowledge Base of European Labour Migration Policies.}
present study. First, it was designed to be representative at regional level and had enough observations to allow a refined analysis of effects for locally defined reference groups. Second, the questionnaire included an entire block of variables related to entrepreneurial behaviour. These comprise detailed information about the activity of the firm (like sector, turn-over, finance, profits, investment, business competitors), retrospective information on the decision to start-up the firm (motivation, sources of finance, sources of information, administrative difficulties) as well as for those who are not entrepreneurs the intentions to become one.

The survey data were matched with regional economic indicators from aggregate statistics. These included information on local economic and labour market conditions like income, unemployment, financial and technical infrastructure, sectoral distribution of economic activities, degree of urbanisation, and foreign direct investments (FDI).

In order to alleviate problems related to the labour force participation for those planning retirement or involved in higher education the paper uses a restricted sample including only individuals between 25 and 50 years old.

4.1 Definition of variables

The analysis used as a main dependent variable the decision of those individuals who were not and are not an entrepreneur to be actively engaged in becoming one. Due to this definition the group of interest is a self-selected one in the sense that it does not include the individual who previously decided to be an entrepreneur and either still are or failed and are not any more entrepreneurs. For the first category the dependent variable is missing since the question is not applicable. This selection will require a correction in the estimation.

As individual controls the analysis used the socio-demographic characteristics available in the data - like age, gender, education, family size, number of children, occupational skills, sector, personal income, household wealth and assets, household income, rural/urban area, size of the locality. Using information from the data set two additional controls were introduced to account for the personal background (Lazear’s Jack-of-all-trade hypothesis) and for the work experience abroad (return migration). The idea behind is that persons who have a larger experience (e.g. due to changes of jobs or sectors) as well as those who have been exposed to work experience abroad (i.e. in a different, more advanced capitalist environment) will have a higher
propensity to take the risks of becoming an entrepreneur. While there is already an indicator of work experience abroad available in the data for the personal background three variables have been merged: numbers of changes in jobs, occupation, and internal mobility for work.

At the household level the wealth variable has been constructed using information on assets and domestic appliances. For the household income remittances received from household members working abroad were also considered in the analysis.

The regional unemployment rate, the regional average income, the cumulated inflows of foreign direct investments over the last 10 years, the population density, and a measurer of infrastructure were used as controls for local effects.

In various specification different types of controls were introduced for contextual effects: like the average skills or education of other in a reference group. A measure of relative deprivation was also constructed to account for the relative wealth and/or income position in the peer-group, which might be important when considering one’s status and costs of failure (stigma). Drawing on Stark (1984), Strak and Taylor (1991) the individual (absolute) income variable of the data set is used to compute a measure of the relative income position with respect to a specific reference group. The relative deprivation measure for individual $i$ is defined as

$$RD^i = \int_{y_i}^{y_m} [1 - F(x)] \, dx,$$  \hspace{1cm} (2)

where

- $y_i$: denotes the income of individual $i$,
- $y_m$: is the maximal income in the individual’s reference group,
- $F(x)$: is the cumulative distribution of income in the reference group.

A second measure of the relative income position of the individual is the ratio of his income to the median income ($y_{med}$) in the corresponding reference group:

$$RI^i = \frac{y_{med}}{y_i}.$$  \hspace{1cm} (3)

Eventually, as an additional measure of income inequality among the members of a reference group a Donaldson-Weymark Gini coefficient$^4$ was computed for each county.

$^4$See Donaldson, Weymark 1980.
One important determinant of entrepreneurship often discussed in the literature is the income premium resulted from being an entrepreneur. Such an income premium is particularly relevant for the purposes of the present study since its role give hints about the presence of non-pecuniary benefits from entrepreneurial activities (like e.g. prestige).

To this end a hypothetical income differential resulted from becoming an entrepreneur was also predicted for each individual in the sample. This was obtained by computing the differential between the actual and the counterfactual wages: i.e. wages predicted for those who had not entrepreneurial activities assuming they did and for those who were not entrepreneurs assuming they were not.

The incomes of entrepreneurs were therefore estimated as

\[ \ln y_{\text{entr}} = \beta_{\text{entr}} X_{\text{entr}} + \varepsilon_{\text{entr}} \] (4)

For those with no entrepreneurial activities (the employed one), their income equation was estimated as:

\[ \ln y_{\text{emp}} = \beta_{\text{emp}} X_{\text{emp}} + \varepsilon_{\text{emp}} \] (5)

For the non-entrepreneurs, with no experience as self-employed the predicted income differential becomes:

\[ \Delta_{\text{inc}} = \frac{e^{\beta_{\text{entr}} X_{\text{emp}}}}{\hat{y}_{\text{emp}}} - \frac{\ln \hat{y}_{\text{entr}}}{e^{\beta_{\text{emp}} X_{\text{ent}}}} \] (6)

while for the entrepreneurs it is:

\[ \Delta_{\text{inc}} = \ln \hat{y}_{\text{entr}} - \hat{\beta}_{\text{emp}} X_{\text{ent}} \] (7)

There is however a double selection problem arising in the estimation. First, like in any other estimation of income equations there is the standard labour force participation selection issue. Second, the entrepreneurs as already mentioned, are as well a non-randomly self-selected group of the population. In order to correct this selectivity bias the income functions were estimated using a double sample selection procedure. The first selection rule concerned the labour force participation:

\[ P_j = (\gamma_j Z_j + u > 0) \] (8)

while the second selection rule for the decision to be an entrepreneur:

\[ S_j = (\delta_j Z_j + v > 0) \] (9)
with the stochastic terms normally distributed

\[ u \sim N(0, 1) \quad (10) \]
\[ v \sim N(0, 1) \quad (11) \]

and allowing the two decisions to be correlated:

\[ \text{corr}(u, v) = \rho \neq 0 \quad (12) \]

In order to identify the endogenous effects two lagged variables were introduced as controls. The first one refers to the ratio of failed entrepreneurs in the reference group. The second is simply the rate of entrepreneurship in the reference group at the present. Since the dependent variable is nascent-entrepreneurship these variables incorporate lagged information - alleviating the reflection problem - because by definition the nascent entrepreneurs can be neither failed nor actual entrepreneurs.

### 4.2 Multiple imputation for non-response

The data set constructed for the present analysis is based on the complete cases for the dependent variables related to migration behaviour. However, since the CURS survey suffers from non-response, i.e. data values are missing although valid values are observable values were imputed for two independent variables, i.e. for log wages, for the cases where all other endogenous variables were observed, and for the non-response cases of the variable capturing the occupational status.

The method used for imputation is based on multiple regressions on the independent variables and assumes that the observations are missing at random (MAR). For log wages OLS regressions are used and the values are imputed only for observations on those who participate in the labour force and declare to be active but do not indicate their income.

Additionally, for the variable on occupational status, values were imputed using multinomial logit regressions. Only the cases with complete observations on the other variables were included, i.e. individuals declaring to actively participate in the labour market and for whom information on wages, sex, age, education, region of origin, employment history after 1989 were available.
4.3 Empirical specification

The latent entrepreneurial decision is specified in an empirical context as:

\[ y_j^* = \beta x_j + \varepsilon_j \]  \hspace{1cm} (13)

whereby \( y_j^* \) is the unobserved expected net present benefit from becoming an entrepreneur, \( x_j \) is a vector of individual and household characteristics, and \( \varepsilon \sim N(0, 1) \). The probit equation is based on the binary outcome \( y_j \), observed in the data set as:

\[ y_j = (y_j^* > 0) \]  \hspace{1cm} (14)

Using the CURS data set, the latent equation on nascent entrepreneurship will be estimated with the functional form given by

\[ y^* = \beta_1 \cdot \text{age} + \beta_2 \cdot \text{age}^2 + \beta_3 \cdot (\sum_{i=1}^{2} \text{GENDER}) + \beta_4 \cdot \text{schooling} \]

\[ + \beta_5 \cdot \text{income} + \beta_6 \cdot (\sum_{j=1}^{2} \text{SECTOR}) + \beta_7 \cdot (\sum_{k=1}^{5} \text{OCCUPATION}) \]

\[ + \beta_8 \cdot (\sum_{l=1}^{42} \text{JOBCHANGES}) + \beta_9 \cdot (\sum_{m=1}^{3} \text{CHILDER}) \]

\[ + \beta_{10} \cdot (\sum_{n=1}^{4} \text{ETHNICITY}) + \beta_{11} \cdot (\sum_{o=1}^{2} \text{URBAN}) \]

\[ + \beta_{12} \cdot (\sum_{p=1}^{2} \text{HHINCOME}) + \beta_{13} \cdot (\sum_{q=1}^{2} \text{RETURN}) \]

\[ + \beta_{14} \cdot \text{incomediff} + \beta_{15} \cdot \text{regunempl} + \beta_{16} \cdot \text{regfdi} \]

\[ + \beta_{17} \cdot (\sum_{s=1}^{2} \text{REMITTANCES}) + \beta_{18} \cdot \text{FAILED ENTREPR} \]

\[ + \beta_{19} \cdot \text{ACTUAL ENTREPR} + k + \varepsilon \]  \hspace{1cm} (15)

Here, \textit{schooling} represents the years of education attained, \textit{SECTOR}, \textit{OCCUPATION}, \textit{REGION}, \textit{CHILD}, \textit{ETHNICITY}, and \textit{URBAN} are sets of dummy variables capturing the ownership form of the employer (private, state, or mixed), the occupation, the presence of children in the household, the ethnic origin and the location in rural or urban areas respectively. \textit{RETURN} is also a dummy variable indicating if the individual has been
working abroad after 1989. *REMITTANCES* is a dummy variable indicating if the household receives remittances from other members currently working abroad.

## 5 Identification of social effects

Manski (1993) criticises the mainstream approach of studying social interactions by regressing the individual outcome on a reference-group’s mean. He shows that this fails to identify the parameter of interactions on outcomes determined at individual level.

Using a linear model Manski(1993) formalises the three hypotheses for explaining social interactions: endogenous effects, contextual effects and correlated effects. If $y$ is a scalar outcome (individual entrepreneurial choices), $z \in \mathbb{R}^K$ individual attributes directly affecting $y$, $u$ a scalar non-observable individual attribute, $x \in \mathbb{R}^J$ attributes of the individual’s reference group, then:

\[
y = \alpha + \beta E(y \mid x) + E(z \mid x)' \gamma + z'\eta + u \tag{16}
\]

\[
u = x'\delta + \epsilon \tag{17}
\]

then $\beta$ captures endogenous effects, $\gamma$ captures contextual effects while $\delta$ captures correlated effects.

\[
E(y \mid x, z) = \alpha + \beta E(y \mid x) + E(z \mid x)' \gamma + z'\eta + x'\delta \tag{18}
\]

\[
E(y \mid x) = \alpha + \beta E(y \mid x) + E(z \mid x)' \gamma + E(z \mid x)' \eta + x'\delta \tag{19}
\]

If $\beta \neq 1$ the solution of (11) is unique:

\[
E(y \mid x) = \frac{\alpha}{1 - \beta} + E(z \mid x)' \frac{\gamma + \eta}{1 - \beta} + E(z \mid x)' \frac{\delta}{1 - \beta} + \frac{x'\lambda}{1 - \beta} \tag{20}
\]

Since $E(y \mid x)$ is a linear function of $[1, E(z \mid x), x]$ it is not possible to distinguish endogenous from exogenous or correlated effects. Glaeser, Scheinkman (2003) discuss some approaches to overcome the identification problem.
In this paper the data allow two alternative measures to be used in order to overcome the identification problem. First, given the large sample size and the fact that the data are representative at county level, the reference group is identified spatially. Second, given that the dependent variable refers to nascent entrepreneurs, the behavioural variables used as controls for endogenous effects are lagged since the decision over nascent-entrepreneurship is by definition a later one in the sequential process of having to decide at an earlier moment to be either an entrepreneur or a failed one.

6 Results

6.1 The empirical puzzle

Before presenting the details of the estimations a bit of intuition on the general empirical puzzle will help understanding the relevance of the question. The initial intriguing relationship in the data was that between the prevalence rate of entrepreneurship (including both the rate of actual entrepreneurs and the ratio of those becoming nascent-entrepreneurs in each county) on the one side and on the other side the estimated average income premia of entrepreneurship per county (computed as described in the section on the empirical specification as the returns to entrepreneurship, i.e. the differential between the hypothetical incomes of non-entrepreneurs if they became one and their actual income). From a theoretical perspective, one would expect a positive relationship between the these income premia and the prevalence rate of entrepreneurship in the corresponding regions. However, as depicted in Figure 2, there is hardly relationship between the two variables and if any it seems to be rather on the opposite, i.e. negative: the higher the income premia the lower the ratio of entrepreneurs.

The second interesting fact observable in the data was the positive relationship between the ration of nascent-entrepreneurs per county and the ration of actual entrepreneurs (see Figure 3). This is already a hint towards the existence of social effects and the presence of non-pecuniary returns from entrepreneurship.

Particularly interesting from the perspective of the present study is the even stronger relationship between nascent entrepreneurship and the per-county rate of failure in entrepreneurial activities. This would be a first indication on endogenous costs of failure (stigma) in the decision making on
entrepreneurial activities.

To test the robustness of these findings the estimated model will have to include a wider range of controls for both regions and individual and to take into account the self-selective nature of entrepreneurial choices.

Across the regions - as Table 3 indicates - there is an important variation in both the ratio of individuals are entrepreneurs and in those who are becoming entrepreneurs (nascent-entrepreneurs). Additionally, Table 4 shows that there is also a substantial variation in economic fundamentals and local labour market conditions across the counties.

As expected, the descriptive statistics depicted in Table 1 show that entrepreneurs tend to be younger, male, be land owners, have a broader personal background, have work experience abroad, and come from richer households with different sources of income (e.g. remittances).
Figure 3: Prevalence rate of nascent-entrepreneurs in Romanian counties

6.2 Determinants of entrepreneurial decisions

The probability of an individual to be a nascent-entrepreneur is estimated in a maximum-likelihood probit model with sample selection. The results for the entrepreneurial decision are presented in Table 5 while the estimated parameters of the selection equation are presented in Table 6.

For most of the individual controls the results are those expected from the existing empirical literature on the determinants of entrepreneurial choices. As the descriptive statistics already suggested, younger males with a level of education above the average and in households with higher incomes and with land ownership display a higher individual probability to be a nascent entrepreneur. Additionally, being in a household which receives remittances from other household members living abroad, as well as being a return migrant themselves, significantly increases the propensity of individuals to be-
come entrepreneurs.

For the purpose of the current study, an interesting result is the estimated parameter of the constructed variable capturing the income premia. In none of the estimated specifications this had a significant effect on nascent entrepreneurs. Moreover, its sign is negative. This is a relevant indicator that non-pecuniary returns from becoming an entrepreneur are likely to significantly affect the decision making process.

The regional controls used in the model are hardly significant. The insignificant effect of the regional unemployment rate suggests that entrepreneurship is not a disguised form of unemployment.

The introduced controls for social interactions, i.e. the rates of failed and of actual entrepreneurship, showed to be highly significant in all the estimated models. This suggest the impact of the aggregate entrepreneurial behaviour in one’s social group on individual decisions.

The individual entrepreneurial choices depend thus not only on economic determinants: peer effect also enter the individual utility function from becoming an entrepreneur.

There is also an intuitive explanation of the results for the selection equation presented in Table 6. Due to the way in which nascent entrepreneurs were defined in this paper and thus extracted from the data set, the positive correlation of the two equations (the decision to be a nascent entrepreneur and the decision to be included in the sample, i.e. not to be already a successful or a failed entrepreneur) could be interpreted as indicating the changing nature of entrepreneurial activities over time as the transition to a market economy progressed. The positive $\rho$ would suggest that those individual who already made a previous decision to become an entrepreneur (and are thus not included in the sample) would be less likely to become at this later moment nascent entrepreneurs assuming that they would have the choice. This is a very plausible conjecture give the widespread rent-seeking behaviour of managers in the earlier stages of transition. Individual who otherwise lack the observable characteristics of entrepreneur did become one in economic circumstances which rewarded such activities in a different way than a market environment.
7 Conclusions

The present paper is the first one to analyse both empirically and theoretically the social dynamics of entrepreneurial choices in a transition economy.

The aim of the paper is to test the presence of social interaction in the decisions over entrepreneurship and to analyse if these social effects could explain the regional variation of entrepreneurial activities.

The theoretical background of the paper is a multiple-equilibria model with endogenous costs of failed entrepreneurial activities (stigma). The results of modelling the entrepreneurial choices in this setting predict the existence of three equilibria: one with low, one with intermediate, and one with high entrepreneurial activity.

In the equilibrium with low entrepreneurial activity there are only few entrepreneurs. Close to the equilibrium there are high rewards for entrepreneurs but still people do not want to become an entrepreneur because the fear of failure is high due to high loss in social prestige in the case of failure. In the equilibrium with high entrepreneurial activity there are many entrepreneurs. Close to the equilibrium there a low rewards for entrepreneurs but still people do try become an entrepreneur because the fear of failure is low. Therefore, the equilibria with high and low entrepreneurial activity are locally stable. This implies that very similar societies and regions might have very different entrepreneurial activities due to a different historical background.

The paper is testing these predictions for the case of a transition economy, Romania, precisely because the transition from a socialist, centralised economy to a capitalist free market system is pertinent to analyse the choices of individuals over entrepreneurial activities.

One important contribution of the empirical part of the paper is the implication of a correlation between individual and aggregated choices with regard to entrepreneurship.

The result show that there are significant social effects on the behaviour of entrepreneurs after controlling for individual and regional characteristics. The presence of negative income premia for entrepreneurs is furthermore suggesting the existence of non-pecuniary returns from entrepreneurship. The fact that peer effects enter the individual utility from entrepreneurial investments has major implication for the future design of measures to promote entrepreneurial activities.
References


Appendix
Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>all sample</th>
<th>entrepreneur</th>
<th>employees</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>age</strong></td>
<td>38.901</td>
<td>39.748</td>
<td>38.825</td>
<td>-.922***</td>
</tr>
<tr>
<td></td>
<td>(.088)</td>
<td>(.264)</td>
<td>(.084)</td>
<td>(.298)</td>
</tr>
<tr>
<td><strong>male</strong></td>
<td>.516</td>
<td>.568</td>
<td>.511</td>
<td>-.056***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.015)</td>
<td>(.004)</td>
<td>(.016)</td>
</tr>
<tr>
<td><strong>log income</strong></td>
<td>1.220</td>
<td>1.514</td>
<td>1.195</td>
<td>-.319***</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.021)</td>
<td>(.005)</td>
<td>(.019)</td>
</tr>
<tr>
<td><strong>occupational change</strong></td>
<td>.422</td>
<td>.615</td>
<td>.405</td>
<td>-.209***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.015)</td>
<td>(.004)</td>
<td>(.015)</td>
</tr>
<tr>
<td><strong>land ownership</strong></td>
<td>.196</td>
<td>.351</td>
<td>.182</td>
<td>-.168***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.017)</td>
<td>(.004)</td>
<td>(.015)</td>
</tr>
<tr>
<td><strong>others’ income in the household</strong></td>
<td>.973</td>
<td>1.200</td>
<td>.952</td>
<td>-.248***</td>
</tr>
<tr>
<td></td>
<td>(.010)</td>
<td>(.043)</td>
<td>(.011)</td>
<td>(.039)</td>
</tr>
<tr>
<td><strong>years of schooling</strong></td>
<td>10.202</td>
<td>11.249</td>
<td>10.158</td>
<td>-1.090***</td>
</tr>
<tr>
<td></td>
<td>(3.477)</td>
<td>(2.657)</td>
<td>(3.501)</td>
<td>(.093)</td>
</tr>
<tr>
<td><strong>hh members abroad</strong></td>
<td>.123</td>
<td>.148</td>
<td>.121</td>
<td>-.026**</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.013)</td>
<td>(.003)</td>
<td>(.012)</td>
</tr>
<tr>
<td><strong>return migrants</strong></td>
<td>.061</td>
<td>.114</td>
<td>.056</td>
<td>-.057***</td>
</tr>
<tr>
<td>(work exp. abroad)</td>
<td>(.002)</td>
<td>(.011)</td>
<td>(.002)</td>
<td>(.009)</td>
</tr>
</tbody>
</table>

Note: The null hypothesis tested refers to mean(employees)-mean(entrepreneurs)=0. *** indicates a .001 and ** a .01 level of significance for rejecting the null hypothesis. Standard errors in parentheses.
Table 2: Descriptives for the dependent variable on nascent entrepreneurs

<table>
<thead>
<tr>
<th></th>
<th>rural</th>
<th>urban</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>switch to entrepreneurship</strong></td>
<td>510</td>
<td>1,394</td>
<td>1,904</td>
</tr>
<tr>
<td></td>
<td>(4.66%)</td>
<td>(9.15%)</td>
<td>(7.27%)</td>
</tr>
<tr>
<td><strong>no switch</strong></td>
<td>9,461</td>
<td>11,387</td>
<td>20,848</td>
</tr>
<tr>
<td></td>
<td>(86.52%)</td>
<td>(74.73%)</td>
<td>(79.65%)</td>
</tr>
<tr>
<td><strong>not applicable</strong></td>
<td>964</td>
<td>2,457</td>
<td>3,421</td>
</tr>
<tr>
<td>(e.g. already entrepreneur)</td>
<td>(8.82%)</td>
<td>(16.12%)</td>
<td>(13.07%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>10,935</td>
<td>15,238</td>
<td>26,173</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(100%)</td>
<td>(100%)</td>
</tr>
</tbody>
</table>

Table 3: Regional variations in entrepreneurship

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
<th>std. dev</th>
<th>minimum</th>
<th>maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>nascent entrepreneurs</td>
<td>.08447</td>
<td>.0849</td>
<td>.03855</td>
<td>.01727</td>
<td>.17102</td>
</tr>
<tr>
<td>current entrepreneurs</td>
<td>.05023</td>
<td>.05118</td>
<td>.01382</td>
<td>.01449</td>
<td>.0821</td>
</tr>
<tr>
<td>failed entrepreneurs</td>
<td>.03899</td>
<td>.03846</td>
<td>.01572</td>
<td>.01447</td>
<td>.07402</td>
</tr>
</tbody>
</table>

Table 4: Regional variation in local conditions

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std. dev</th>
<th>minimum</th>
<th>maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>unemployment</td>
<td>10.99787</td>
<td>3.173501</td>
<td>4.5</td>
<td>16.7</td>
</tr>
<tr>
<td>income</td>
<td>2.03553</td>
<td>.2614472</td>
<td>1.673601</td>
<td>2.77453</td>
</tr>
<tr>
<td>FDI</td>
<td>127.9263</td>
<td>406.9828</td>
<td>1.5361</td>
<td>2424.536</td>
</tr>
<tr>
<td>population density</td>
<td>90.93584</td>
<td>31.71396</td>
<td>30.9</td>
<td>181.4</td>
</tr>
<tr>
<td>infrastructure</td>
<td>47.64823</td>
<td>19.25336</td>
<td>8.4</td>
<td>115</td>
</tr>
<tr>
<td>opportunity entrepr.</td>
<td>.0210522</td>
<td>.0100888</td>
<td>.0075188</td>
<td>.0708215</td>
</tr>
</tbody>
</table>
Table 5: The nascent entrepreneur decision
Probit model with sample selection for nascent entrepreneurs

<table>
<thead>
<tr>
<th>variable</th>
<th>coeff.</th>
<th>robust std. error</th>
<th>dy/dx</th>
<th>std. error X</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>.0161</td>
<td>(.0179)</td>
<td>.0014</td>
<td>(.0016)</td>
</tr>
<tr>
<td>age²/100</td>
<td>-.0498</td>
<td>(.0214)**</td>
<td>-.0044</td>
<td>(.0021)**</td>
</tr>
<tr>
<td>male</td>
<td>.2284</td>
<td>(.0548)**</td>
<td>.02120</td>
<td>(.0052)**</td>
</tr>
<tr>
<td>others' income</td>
<td>.0233</td>
<td>(.0049)**</td>
<td>.00210</td>
<td>(.0004)**</td>
</tr>
<tr>
<td>income differential</td>
<td>-.1538</td>
<td>(.1231)</td>
<td>-.01383</td>
<td>(.0110)</td>
</tr>
<tr>
<td>other hh business</td>
<td>.0954</td>
<td>(.0337)**</td>
<td>.00858</td>
<td>(.0030)**</td>
</tr>
<tr>
<td>medium</td>
<td>.1327</td>
<td>(.0635)**</td>
<td>.01238</td>
<td>(.0060)**</td>
</tr>
<tr>
<td>vocational</td>
<td>.2843</td>
<td>(.0992)**</td>
<td>.03083</td>
<td>(.0124)**</td>
</tr>
<tr>
<td>high</td>
<td>.3677</td>
<td>(.1081)**</td>
<td>.04128</td>
<td>(.0143)**</td>
</tr>
<tr>
<td>child</td>
<td>-.0392</td>
<td>(.0302)</td>
<td>-.00353</td>
<td>(.0027)</td>
</tr>
<tr>
<td>land owner</td>
<td>.1848</td>
<td>(.0483)**</td>
<td>.01821</td>
<td>(.0052)**</td>
</tr>
<tr>
<td>return migrant</td>
<td>.3455</td>
<td>(.0940)**</td>
<td>.03107</td>
<td>(.0087)**</td>
</tr>
<tr>
<td>remittances</td>
<td>.3795</td>
<td>(.1571)**</td>
<td>.04638</td>
<td>(.0251)*</td>
</tr>
<tr>
<td>reg. unemployment</td>
<td>.0077</td>
<td>(.0103)</td>
<td>.00069</td>
<td>(.0009)</td>
</tr>
<tr>
<td>reg. income</td>
<td>.2516</td>
<td>(.1788)</td>
<td>.02262</td>
<td>(.0162)</td>
</tr>
<tr>
<td>reg FDI</td>
<td>.0011</td>
<td>(.0005)**</td>
<td>.00010</td>
<td>(.0000)**</td>
</tr>
<tr>
<td>population density</td>
<td>-.0035</td>
<td>(.0001)**</td>
<td>-.00003</td>
<td>(.0000)**</td>
</tr>
<tr>
<td>infrastructure</td>
<td>-.0003</td>
<td>(.0015)</td>
<td>-.00002</td>
<td>(.0001)</td>
</tr>
<tr>
<td>failed entrepreneurship</td>
<td>4.7659</td>
<td>(1.909)**</td>
<td>4.2853</td>
<td>(.169)**</td>
</tr>
<tr>
<td>actual entrepreneurship</td>
<td>4.5211</td>
<td>(2.479)**</td>
<td>4.0652</td>
<td>(.2301)**</td>
</tr>
<tr>
<td>const.</td>
<td>-3.0160</td>
<td>(.5168)**</td>
<td>.42853</td>
<td>(.169)**</td>
</tr>
</tbody>
</table>

dy/dx is for discrete change of dummy variable from 0 to 1
Table 6: The selection equation

<table>
<thead>
<tr>
<th>variable</th>
<th>coeff.</th>
<th>robust std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.01367</td>
<td>(0.0024)*****</td>
</tr>
<tr>
<td>male</td>
<td>-0.0879661</td>
<td>(0.0314)*****</td>
</tr>
<tr>
<td>father’s occupation (medium)</td>
<td>-0.0737889</td>
<td>(0.0452)</td>
</tr>
<tr>
<td>father’s occupation (high)</td>
<td>-0.2774505</td>
<td>(0.0565)*****</td>
</tr>
<tr>
<td>occupation 1989 (medium)</td>
<td>0.0029165</td>
<td>(0.0420)**</td>
</tr>
<tr>
<td>occupation 1989 (high)</td>
<td>-0.2255026</td>
<td>(0.0564)*****</td>
</tr>
<tr>
<td>land owner</td>
<td>-0.1925602</td>
<td>(0.0455)*****</td>
</tr>
<tr>
<td>home owner</td>
<td>0.0748337</td>
<td>(0.0438)*</td>
</tr>
<tr>
<td>wealth (assets)</td>
<td>-0.4460333</td>
<td>(0.0450)*****</td>
</tr>
<tr>
<td>constant</td>
<td>0.7203743</td>
<td>(0.1107)*****</td>
</tr>
</tbody>
</table>

/athrho                               | 0.6953073  | (0.1761)*****    |

rho                                   | 0.5013807  | (0.1124)         |

Wald test of indep. eqns. (rho = 0): chi2(1) = 15.58