

Gang, Ira N.; Co, Catherine Y.; Yun, Myeong-Su

**Working Paper**

## Switching Models with Self-Selection: Self-Employment in Hungary

Working Paper, No. 1999-12

**Provided in Cooperation with:**

Department of Economics, Rutgers University

*Suggested Citation:* Gang, Ira N.; Co, Catherine Y.; Yun, Myeong-Su (1999) : Switching Models with Self-Selection: Self-Employment in Hungary, Working Paper, No. 1999-12, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/94337>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Switching Models with Self-Selection: Self-Employment in Hungary

by

Catherine Y. Co  
Department of Economics  
University of Central Florida

Ira N. Gang  
Department of Economics  
Rutgers University

Myeong-Su Yun  
Department of Economics  
Rutgers University

August 24, 1999 (4:24PM)

Correspondence to:

Ira N. Gang  
Department of Economics  
Rutgers University  
75 Hamilton St  
New Brunswick NJ 08901-1248 USA  
phone: (+1 732) 932-7405 fax: (+1 732) 932-7416  
email: gang@economics.rutgers.edu

JEL codes: C34, J31, P23

Keywords: self-selection, switching, decomposition, self-employment, transition

## Abstract:

We estimate the determinants of earnings for both the self-employed and wage/salary sectors in an economy undergoing transition from socialism to greater market orientation. We adopt a (full information) MLE methodology in addition to Heckman's two-step method, while taking both participation and self-employment decisions into account. We use the Hungarian Household Panel Survey for 1994. We find that the return to characteristics are not significantly different between self-employed and wage/salary sectors. However, when we also account for the dispersion of earnings, the structure of the earnings between the self- and wage/salary employed are different.

\* We thank seminar participants at IZA, Bonn for their comments.

## I. Introduction

Measures to promote self-employment are actively considered by policy makers. Self-employment is seen as an alternative to wage/salary employment and thereby broadens the choices facing both the potential entrant to the labor market and the underemployed. A natural question that arises is whether workers with the same productivity traits receive equal remuneration in self- and wage/salary employment. Is there a self-employment remuneration advantage?

Previous researchers generally find that the self-employed earn more than other workers in developing countries (Sumner (1981), Blau (1985, 1986), Vijverberg (1986)). We investigate whether this is true in a transition economies by studying self- versus wage/salary employment in Hungary, one of several countries undergoing transition into a market economy. Moreover, earlier studies usually correct for only one form of selection bias, whether those entering self-employment are self-selected. We allow for selection into labor force participation or not, as well as a switching regression as to whether one is self- or wage/salary employed. We implement a natural method for using maximum likelihood estimation for this type of situation. Indeed, we believe the method we implement here should find application in a wide range of labor market studies.

Self-selection into self-employment versus wage/salary and into participation versus nonparticipation may be an important consideration in transition economies. This may be especially true in Hungary. Early on, Hungary followed policies that encouraged self-employment, as part of their move during the 1970s and 1980s toward market-oriented socialism. Hungary has promoted self-employment during transition as a way to attack unemployment problems (O'Leary (1999)). The role and performance of the self-employed in Hungary may serve as an example for other transition economies.

As early as 1964, Hungary committed itself to the establishment of a market-oriented socialist economy, formalizing this in 1968 with the adoption of the *New Economic Mechanism*.

An environment conducive to the development of private enterprises was stimulated by the 1988 Company Act that allowed the formation of modern forms of business associations and provided basic guarantees for private and foreign investors (OECD, 1991, p. 79). In 1990, Hungary created the State Property Agency. This agency was given the mandate to privatize state-owned properties. These initiatives led to the growth in the number of new private business creations with Hungary experiencing a 50% growth in the number of entrepreneurs in 1990 (Sziraczki, 1993). The number of registered non-agricultural entrepreneurs rose from approximately 400,000 to more than 700,000 from 1990-94 (Gabor, 1994). By 1994, self-employment accounts for 20% of non-agricultural employment in Hungary (Gabor, 1994).<sup>1</sup>

In transition economies, there are at least two sources of the self-employed. First, these are the same individuals who were already operating small-scale private businesses before 1989 (Frey and Timar, 1993, p. 184). In the pre-transition era this was referred to as the second economy. Sziraczki (1993) estimates that in the middle of the 1980s, about three-fourths of all Hungarian households had additional income from activities in the second economy, gaining entrepreneurial experience. When new policies were adopted in the late 1980s, the economic activities of these individuals were fully legalized and formally recognized. The second source of the self-employed are "displaced" workers.<sup>2</sup> Privatization usually involves labor displacement as enterprises undergo restructuring. Those displaced, having no source income, can turn to self-employment. In 1991, a scheme was adopted whereby the unemployed can invest their unemployment benefits and undergo training to start their own business.<sup>3</sup> In our data, we cannot distinguish between these two

---

<sup>1</sup> The comparable figure for OECD countries is 11%, 15% for North Africa and the Middle East, 26% for Latin America and about 27% for Asian countries (See Gabor, 1994, p. 339).

<sup>2</sup> Unemployment rate in the second quarter of 1994 was 12.4%, compared to 0.7% in the second quarter of 1990 (see OECD, 1995, p. 51).

<sup>3</sup> O'Leary (1999) reviews and analyzes self-employment programs in Hungary during transition.

sources of self-employment.

In Section II we develop our econometric model. The data is discussed in Section III. Results of our analysis are presented in Section IV, and conclusions in Section V.

## II. Econometric Models

Papers that estimate the earnings equations of wage/salary earners versus the self-employed typically ignore non-labor force participants and estimate the earnings equations using the switching regression model (for example, see Rees and Shah (1986), Gill (1988), and Yuengert (1995).<sup>4</sup> The switching regression accounts for the self-selection that occurs in the wage/salary versus self-employment decision. We argue that ignoring non-labor force participants will lead to biased estimates of the earnings regressions. One of our contributions is to estimate the earnings equation accounting for two self-selection decisions. We take advantage of the well-known property that switching regression and self-selectivity models have much in common, and can be estimated by similar methods.<sup>5</sup>

Individuals choose either self- or wage/salary employment, depending upon from which he/she receives the highest utility. That is, individuals choose self- or wage/salary employment according to the following index,

$$S^* = Q\theta + v, \quad (1)$$

where  $\theta$  is vector of coefficients and  $v$  is a stochastic term. Individuals choose self-employment if

---

<sup>4</sup> To be precise, they use an endogenous switching regression model with known sample separation. See Maddala (1986) for details.

<sup>5</sup> Maximum likelihood estimation is usually used for the switching model (especially for models with unknown sample separation), while Heckman's two-step method is widely used for the common self-selection bias correction model. However, both models can be estimated using either method. Srivastava and Rao (1990) show the application of both Heckman's two-step and maximum likelihood estimation methods to switching models (including the case with unknown sample separation).

$S^*$  is positive; otherwise, they choose wage/salary employment. Though  $S^*$  is unobservable, we can observe individuals' choice, a dichotomous variable  $S$ , ( $S = 1$  if  $S^* > 0$ , and  $S = 0$  otherwise).

The earnings function for each choice is,

$$Y_j = \mathbf{X}\beta_j + e_j, \quad (2)$$

where  $j$  is 1 (self-employment) or 0 (wage/salary job),  $Y$  is the natural-log of monthly earnings, and  $e$  is a stochastic term.

The estimates of equation (2) from ordinary least square (OLS) may be inconsistent due to switching between self- and wage/salary employment. We can use either Heckman's two-step method (see Heckman (1979)) or maximum likelihood estimation (MLE) to obtain consistent estimates.

Following the standard model, we assume the stochastic terms ( $e_j, \nu$ ) follow a joint normal distribution with mean zero and the following variance-covariance matrix:

$$\Sigma_j = \begin{bmatrix} \sigma_{e_j}^2 & \sigma_{e_j\nu} \\ & \sigma_\nu^2 \end{bmatrix},$$

where  $\sigma_\nu^2$  is normalized to 1.<sup>6</sup>

Heckman's two-step method estimates the switching equation (1) by probit, and defines the selection bias correction term,  $\lambda_{1j}$ , using the estimates from probit.  $\lambda_{1j}$  is defined as  $\phi(-\tilde{\mathbf{S}}) \cdot \Phi(\tilde{\mathbf{S}})^{-1}$  when  $j$  is 1, and  $-\phi(-\tilde{\mathbf{S}}) \cdot \Phi(-\tilde{\mathbf{S}})^{-1}$  when  $j$  is 0, where  $\tilde{\mathbf{S}} = \mathbf{Q}\theta/\sigma_\nu$ , and  $\phi$  and  $\Phi$  are standard normal probability and distribution density function, respectively. At the second step, OLS is used to estimate the earnings equation (2) which now also includes  $\lambda_{1j}$ .

---

<sup>6</sup> It is usual to present the covariance matrix of ( $e_1, e_0, \nu$ ). Since the covariance of  $e_1$  and  $e_0$  is not identified in this kind of model, for simplicity we split the covariance matrix into the two covariance matrices of ( $e_1, \nu$ ) and ( $e_0, \nu$ ). See Koop and Poirier (1997) and Vijverberg (1993) for a discussion of the identification issue for the covariance of ( $e_1, e_0$ ).

On the other hand, we can maximize the following likelihood,<sup>7</sup>

$$L = \prod \left[ \Phi \left( \frac{Q\theta + \mu_{vle_1}}{\sigma_{vle_1}} \right) \cdot \frac{1}{\sigma_{e_1}} \cdot \phi \left( \frac{e_1}{\sigma_{e_1}} \right) \right]^s \cdot \left[ \Phi \left( \frac{-Q\theta - \mu_{vle_0}}{\sigma_{vle_0}} \right) \cdot \frac{1}{\sigma_{e_0}} \cdot \phi \left( \frac{e_0}{\sigma_{e_0}} \right) \right]^{(1-s)} \quad (3)$$

where  $\mu_{vle_j} = \rho_{e_jv} \sigma_v \sigma_{e_j}^{-1} e_j$ ,  $\sigma_{vle_j} = \sigma_v (1 - \rho_{e_jv})^{0.5}$ , and  $\rho_{e_jv} = \sigma_{e_jv} (\sigma_{e_j} \sigma_v)^{-1}$ ,  $j = 1$  or  $0$ .

The above equations establish the earnings of individuals taking into account the choice of type of employment. Samples employed for estimating the above equations exclude people who are not working. However, the self-selection of workers into working or not causes a bias in the estimates of the earnings equation.<sup>8</sup> The inconsistency of OLS estimates occurs not only due to the self-employment choice but also due to the working selection. To obtain consistent estimates of coefficients in the wage/salary equations we must take account of both the selection into working or not and the choice of self-employment versus wage/salary employment. We picture the selection process as having two steps: first, individuals determine whether to work or not; second, individuals who have decided to work choose self- or wage/salary employment.

To address the decision to work or not, we introduce a second index function,

$$P^* = Z\gamma + u, \quad (4)$$

where  $\gamma$  is vector of coefficients and  $u$  is stochastic term. Individuals choose to work if  $P^*$  is positive; otherwise they do not work. We observe the dichotomous variable  $P$ , which has a value

---

<sup>7</sup> Equation (3) is the functional expression of the following,

$$L = \prod [Pr(S=1|Y_1) \cdot Pr(Y_1)]^s \cdot [Pr(S=0|Y_0) \cdot Pr(Y_0)]^{1-s}.$$

<sup>8</sup> Correcting for the bias that arises due to workers' self-selecting themselves into work is standard for women, but not for men. However, the men in our sample have a working rate of 54%, while the women have a working rate of 48%. This is very low for prime-aged males, so considering the work decision for males here is appropriate.

of 1 if  $P^* > 0$ , and zero otherwise.

Similar to the standard switching regression model or self-selection bias model, we can apply Heckman's two-step method or MLE method to estimate the effects of participation and self-employment choices on earnings.

We assume the stochastic terms ( $e_j, v, u$ ) follow a joint normal distribution with mean zero and the following variance-covariance matrix:

$$\Sigma'_j = \begin{bmatrix} \sigma_{e_j}^2 & \sigma_{e_j v} & \sigma_{e_j u} \\ & \sigma_v^2 & \sigma_{vu} \\ & & \sigma_u^2 \end{bmatrix},$$

where  $\sigma_v^2$  and  $\sigma_u^2$  are normalized to 1.

Heckman's two-step method first estimates the switching equation (1) and participation selection equation (4) by bivariate probit, and defines two selection bias correction terms ( $\lambda_{1j}$  and  $\lambda_{2j}$ ) using the estimates from the bivariate probit (for details, see Fische, Trost and Lurie (1981), Ham (1982), and Tunali (1986)). In the second step, OLS is used to estimate the earnings equation (2) which include  $\lambda_{1j}$  and  $\lambda_{2j}$ . The selection correction terms are defined as following,

$$\lambda_{11} = \phi(-\tilde{S}) \cdot \left[ 1 - \Phi \left( \frac{-\tilde{P} - \rho_{vu} \tilde{S}}{(1 - \rho_{vu})^{0.5}} \right) \right] \cdot \Psi(\tilde{S}, \tilde{P}, \rho_{vu})^{-1}, \text{ and}$$

$$\lambda_{21} = \phi(-\tilde{P}) \cdot \left[ 1 - \Phi \left( \frac{-\tilde{S} - \rho_{vu} \tilde{P}}{(1 - \rho_{vu})^{0.5}} \right) \right] \cdot \Psi(\tilde{S}, \tilde{P}, \rho_{vu})^{-1}, \text{ for } j = 1,$$

and

$$\lambda_{10} = -\phi(-\tilde{S}) \cdot \left[ 1 - \Phi \left( \frac{-\tilde{P} - \rho_{vu} \tilde{S}}{(1 - \rho_{vu})^{0.5}} \right) \right] \cdot \Psi(-\tilde{S}, \tilde{P}, -\rho_{vu})^{-1}, \text{ and}$$



$$\lambda_{20} = \phi(-\tilde{P}) \cdot \Phi\left(\frac{-\tilde{S} - \rho_{vu} \cdot \tilde{P}}{(1 - \rho_{vu})^{0.5}}\right) \cdot \Psi(-\tilde{S}, \tilde{P}, -\rho_{vu})^{-1}, \text{ for } j = 0,$$

where  $\tilde{P} = Z\gamma/\sigma_u$ ,  $\tilde{S} = Q\theta/\sigma_v$ , and  $\phi$ ,  $\Phi$ , and  $\Psi$  are the standard univariate probability and distribution, and standard bivariate normal distribution function, respectively.

Though Heckman's two-step method is used to obtain 'consistent' estimators in the presence of double selection in previous papers, Heckman's two-step method becomes cumbersome when the number of selection rules is more than one. This is because the formulae for the computation of the self-selection correction terms ( $\lambda$ 's) become complicated, as shown above, and computing corrected standard errors requires nontrivial programming.<sup>9</sup>

We adopt a (full information) MLE method in addition to Heckman's two-step method for our study. MLE is an attractive method to estimate earnings equation and two selection functions jointly (e.g., see Blank (1990), and Co, Gang, Yun (1999)).<sup>10</sup> The procedure accounts for the endogeneity of the participation and self-employment decisions with earnings, which is ignored by OLS (see Heckman (1978) and Moffitt (1983), p. 1030). The obtained estimators are not only consistent, but also have other desirable properties of MLE (they are asymptotically efficient and normally distributed). MLE becomes easy to implement when the stochastic terms are assumed to follow a multivariate normal distribution.

---

<sup>9</sup> The burden of computation can be relieved by assuming that the two selections are not correlated (Fishe, Trost and Lurie (1981)). However, this is often too strong an assumption.

<sup>10</sup> In this context MLE has only been infrequently employed. We suspect this is because: 1) researchers have gotten used to the two-step procedure and like to see the  $\lambda$ 's included and interpreted; and 2) the likelihood function typically varies from specification to specification, many researchers feel more comfortable with a standard approach and form. The popularity of Heckman's two-step method can also be attributed to its availability in computer packages. Recent developments in optimization programs enable us to more easily estimate using MLE. Here we offer an MLE implementation that is tractable and easily reproduced in problems with a similar structure.

The likelihood function is,<sup>11</sup>

$$L = \prod \left[ \Psi \left( \frac{Q\theta + \mu_{v|e_1}}{\sigma_{v|e_1}}, \frac{Z\gamma + \mu_{u|e_1}}{\sigma_{u|e_1}}, \rho_{vu|e_1} \right) \cdot \frac{1}{\sigma_{e_1}} \cdot \phi \left( \frac{e_1}{\sigma_{e_1}} \right) \right]^{S \cdot P} \cdot \left[ \Psi \left( \frac{-Q\theta - \mu_{v|e_0}}{\sigma_{v|e_0}}, \frac{Z\gamma + \mu_{u|e_0}}{\sigma_{u|e_0}}, -\rho_{vu|e_0} \right) \cdot \frac{1}{\sigma_{e_0}} \cdot \phi \left( \frac{e_0}{\sigma_{e_0}} \right) \right]^{(1-S) \cdot P} \cdot \left[ \Phi \left( -\frac{Z\gamma}{\sigma_u} \right) \right]^{1-P} \quad (5)$$

where  $\rho_{e_j k} = \sigma_{e_j k} (\sigma_{e_j} \sigma_k)^{-1}$ ,  $\mu_{kle_j} = \rho_{e_j k} \sigma_k \sigma_{e_j}^{-1} e_j$ ,  $\sigma_{kle_j} = \sigma_k (1 - \rho_{e_j k})^{0.5}$ ,  $\sigma_{vule_j} = \sigma_v \sigma_u (\rho_{vu} - \rho_{e_j v} \rho_{e_j u})$ , and  $\rho_{vule_j} = \sigma_{vule_j} (\sigma_{v|e_j} \sigma_{u|e_j})^{-1}$  for  $k = v$  or  $u$ , and  $j = 1$  or  $0$ .

The likelihood function (5) shows the contribution of individuals who are working and self-employed ( $P=1, S=1$ ), individuals who are working and wage/salary earners ( $P=1, S=0$ ), and individuals who are not working ( $P=0$ ), respectively. By maximizing the likelihood function, we obtain estimators of the index functions (working and self-employment decision functions,  $\gamma$  and  $\theta$ ), the earnings function ( $\beta_j$ ), and variance and correlation coefficients. The MLE is implemented using the SAS NonLinear Programming procedure (SAS Institute, 1997).

### III. Data

We use the Hungarian Household Panel Survey (HHPS), a unique data set collected by the Social Research Informatics Centre, Budapest University of Economics. The first wave of the survey was drawn in 1992 (see Sik (1995) for a description). We draw our sample from the 1994 wave of the survey, supplemented with data from the 1993 wave. Our sample consists of those individuals who are in their working “life,” that is, those between 18 and 65 years old in 1994. Out of 3145

---

<sup>11</sup> Equation (5) is the functional expression of followings;

$$L = \prod \left[ Pr(S=1, P=1|Y_1) \cdot Pr(Y_1) \right]^{S \cdot P} \cdot \left[ Pr(S=0, P=1|Y_0) \cdot Pr(Y_0) \right]^{(1-S) \cdot P} \cdot \left[ Pr(P=0) \right]^{1-P}.$$

individuals, 121 individuals were identified as self-employed.

Table 1 contains the means of the variables used in the analysis, for all observations and for those who are working. For each variable we test the null hypothesis that the mean for those who are self-employed is equal to the mean for those who are wage/salary earners. The earnings variable is monthly earnings from a person's main job (natural log of monthly earnings in forints). Using all observations (men and women combined) there is evidence at the 5% significance level that the self-employed earn significantly more than the wage/salary earners. The significance holds at the 1% level for men; but the difference for women are not statistically significant. A significantly smaller percentage of the self-employed are women. Head of household is equal to one for those who are heads of household. There is evidence that a larger percentage of men who are self-employed are heads of household. The gender variable, female, has the values "1" for women and "0" for men. No other variables show a significant difference between the self-employed and wage/salary earners.

**Table 1 near here**

The other variables used are age, experience, education, marital status, the number of children less than six in the family and current residence of the individual. Though not significant, men's age, number of years of education and work experience (actual years working) are larger for men who are self-employed.<sup>12</sup> The opposite holds for women's age and experience. Women who are wage/salary earners are somewhat older and have more experience. Both men and women wage/salary earners have more children less than six, again the difference is not statistically significant. Finally, relatively more women wage/salary earners are living in Budapest and relatively more men who are self-employed are living in Budapest.

The determinants of participation (the decision to work or not) includes age, number of years

---

<sup>12</sup> Unfortunately, the data do not include information on the number of years of self-employed experience. Experience is an aggregate of wage/salary and self-employment experience.

of formal education, marital status, head of household and the number of children below six. Age and its square term are included in the labor force participation equation to test the notion that the probability of work increases with age up to a point, then it starts to decline. Investments in formal education are made with the expectation of higher earnings. This implies that the probability of work rises with the number of years of schooling. The number of children less than six is interacted with female. The number of young children is expected to affect women's decision to work or not.

The self-employment versus wage/salary employment decision includes age, head of household, educational attainment and female. Age and its square term are included. Younger individuals try riskier occupations first (Johnson (1978) and Miller (1984)). This increases the probability of self-employment. However, as a person ages, risk aversion increases and, hence, the probability of self-employment decreases with age. Head of household is included to test the notion that heads of household are more responsible, hence have the necessary abilities to run their own businesses. On the other hand, being the primary source of earnings may make an individual more risk averse and take on a wage-earning job. Finally, female is included to test the notion that men are more likely to be self-employed than women.

Earnings determinants include personal and human capital characteristics. Whether an individual is living in Budapest or not is included to control for earnings and cost-of-living differentials across locations. Education and experience (and its square term) are used as proxies for human capital. Education is expected to have a positive effect on earnings; while the effect of experience is assumed to rise and then fall. Finally, we also control for gender earnings differentials.

#### **IV. Results**

For our analysis we are estimating three models using three estimation methods: 1) a simple earnings equation model (estimated with OLS); 2) a single switching model in which we account

for the choice of self-employment versus wage/salary earners in estimating earnings (estimated by Heckman's two-step method and by MLE); 3) a switching and selection model in which we account for both the choice of self-employment versus wage/salary earners and the choice between working or not in estimating earnings (estimated by Heckman's two-step method and by MLE). In Section A (below) we discuss the determinants of the choice of self-employment versus wage/salary earners and the choice between working or not in estimating earnings in our latter two models. In Section B we discuss the determinants of earnings of the self-employed and wage/salary earners from our three models.

### **A. Determinants of Self-Employment and Participation**

We first consider the results for the self-employment (switching) and participation (selection) equations. The first two columns of Table 2 present the estimates for the single switching equation model (wage/salary versus self-employment decision) from the Heckman's two step and MLE methods, ignoring the labor force participation decisions. Columns 3-6 of the table contain estimates for both switching (wage/salary versus self-employment) and selection (working or not) equation models from both estimation methods.

#### **Table 2 near here**

We can examine coefficient estimates for self-employment choice across our two models (switching and switching-selection) by comparing column 1 to column 4 and column 2 to column 6 in Table 2. Generally, the estimates using these two models are somewhat different, but the significance of the estimates does not change except for the constant term. The coefficient estimates for the head of household are significant and similar across models. The results indicate that the probability of self-employment is larger for individuals who are heads of household. There is no evidence that men are significantly more likely to be self-employed than women.

The probability of entering self-employment is not significantly affected by age. This

finding is consistent with the findings of Evans and Leighton (1990).<sup>13</sup> We explain the insignificant coefficient for age as follows: younger individuals are expected to be more receptive to risk, however, in the context of a transition economy where the financial system is not well developed, these individuals would have less resources to start their own businesses. On the other hand, older individuals may have accumulated the necessary resources to start their own businesses but are less willing to take risks. The closer an individual is to retirement, the less probability he would risk his savings.

Education is insignificant in the choice between self- and wage/salary employment. We expect self-employment to increase with education. However, a person's employability also increases with education. To prospective employers, educational attainment is usually a signal of a person's potential and reliability as a worker. Those with more education may choose wage/salary employment because of the larger potential earnings. These two effects may be canceling each other out giving rise to an insignificant education variable.

The results for the labor participation equation are consistent with results from previous studies. The coefficient on age is positive and the coefficient on age-squared is negative. Investments are made in formal education with the expectation of increasing one's earnings, and we find the probability of work increases with education. The probability of participation in the labor force is significantly higher for married individuals. The number of children less than six significantly decreases women's participation in the labor force.

### **B. Determinants of Earnings**

Tables 3 and 4 present the coefficient estimates from the earnings equations for the

---

<sup>13</sup> The authors use a longitudinal data set. Using U.S. data, they find that the probability of starting a business is independent of age. They explain this as follows: it takes time for people to discover opportunities; on the other hand, older individuals are less likely to take the chance accorded by the opportunity.

wage/salary earners and the self-employed, respectively. Column 1 presents the OLS estimates. The estimates in columns 2 and 3 ignore the labor force participation decision (single switching model) using Heckman's two-step and MLE methods. Columns 4 and 5 contain estimates that take both participation and the self- versus wage/salary employment decisions into account (switching and selection model).

### **Table 3 and 4 near here**

For wage/salary earners, the coefficient estimates of the earnings equation are robust across the single earnings equation model, the single switching model and the switching and selection model. This is not true for the self-employed when MLE is used for our single switching and switching and selection models. Once participation choice is considered for the self-employed (i.e., in the switching and selection model), the estimates of the education and experience parameters become significant at 5% level and larger than those of single switching model.

In comparing the results across wage/salary earners and the self-employed, we focus on the results in columns 4-5, in particular, column 5 where we use MLE. Educational attainment is statistically significant in earnings equations of both the wage/salary earners and the self-employed at the 10% and 5% level, respectively. For each additional year of education, wage/salary earners earn about 8% more while the self-employed earn about 9% more.

Experience has a positive, and experience squared term has a negative, effect on earnings. Interestingly, log-earnings for the wage/salary earners peaks at 30 years of experience while for the self-employed, it peaks at 19 years of experience. Not only do the self-employed on average earn more (see Table 1), but they reach their highest earnings level earlier. This is because individuals on wage/salary employment need to move up the ranks. Regarding the coefficient estimates for the other variables, there is significant earnings differential across locations (Budapest versus other areas) for wage/salary earners but not for the self-employed. Wage/salary earners in Budapest earn

a premium of about 17%. The same situation applies to female. Gender contributes to earnings differential only for the wage/salary earners. Female wage/salary earners earn about 23% less than male wage/salary earners.<sup>14</sup>

Are the returns to attributes the same for the wage/salary earners and the self-employed? We have tested whether there is a structural difference between the earnings of the self-employed and the earnings of wage/salary workers using the likelihood ratio (LR) test. The null hypothesis is  $\beta_0 = \beta_1$ , where  $\beta_0$  and  $\beta_1$  are coefficients of earnings equation parameters for the wage/salary earners and the self-employed, respectively. From the results of the LR test, we cannot reject the null hypothesis, because the critical values ( $-2\lambda$ ) are 5.523 and 6.927 and their probabilities from  $\chi^2$  distribution with 6 degree of freedom are 0.521 and 0.673 for single switching and switching and selection models, respectively. The results indicate that the return to the characteristics are not different in both wage/salary and self-employed sectors.

The results of the LR test might indicate that pecuniary rewards to attributes of the wage/salary earners and the self-employed do not play large role in job choice. Preferences toward different jobs may be more important than the remuneration of a job. The unavailability of the wage/salary jobs to some group of labor market participants may also be the reason for choosing self-employment. Whatever the cause of choosing self-employment, we can expect that the earnings distribution of the self-employed will be less egalitarian than that of wage/salary earners. Hence the structure of earnings equations will be different if we consider the earnings distribution. Using the standard deviation of log-earnings as a measure of earnings inequality, we implement another LR test whose null hypothesis is  $\sigma_0 = \sigma_1$  in addition to  $\beta_0 = \beta_1$ . In other words, we are testing whether the return to human capital is the same ( $\beta_0 = \beta_1$ ) and whether the dispersion of earnings is the same

---

<sup>14</sup>These percentages are calculated using  $\exp[\beta \cdot 0.5 V(\beta)] - 1$ , where  $\beta$  is the estimated coefficient and  $V(\beta)$  is the variance of  $\beta$  (see Kennedy's (1981)).



( $\sigma_0 = \sigma_1$ ). From the results of the LR test, we reject the null hypothesis, because the probabilities from  $\chi^2$  distribution with 7 degree of freedom for the critical values ( 64.658 for the single switching model and 57.337 for the switching and selection model) are one for both models.

## V. Conclusion

We estimate the determinants of earnings for both the self-employed and wage/salary sectors in an economy undergoing transition from socialism to greater market orientation. We adopt a (full information) MLE methodology in addition to Heckman's two-step method, while taking both participation and self-employment decisions into account. Our MLE methodology has application in a wide range of labor market studies. We use the Hungarian Household Panel Survey from 1994, with supplemental information from 1993.

Two selection issues arise: First, those who choose self-employment over wage/salary employment may be self-selected. For example, they may have done better (or worse) during the transformation phase regardless of whether or not they had chosen self-employment. In addition, we face the standard labor force participation (usually, working or not working) selection issue as in any earnings estimation problem. We handle these two issues and the earnings equation estimation jointly, by implementing a tractable and easily reproducible maximum likelihood estimation.

Most labor market studies exclude the self-employed from their analyses. However, if there is self-selection, this exclusion may lead to inconsistent estimates. We find evidence that there is significant negative correlation between the self-employment equation and the earnings equation for wage/salary earners ( $\rho_{e_0v}$ ). This is evidence that the person who has more unobserved characteristics suitable for the wage/salary job, is in a wage/salary job. We also find evidence of significant positive correlation between the participation equation and the earnings equation for the

self-employed ( $\rho_{e_1u}$ ). We find that the return to characteristics are not significantly different between self-employed and wage/salary sectors. However, when we also account for the dispersion of earnings, the structure of the earnings between the self- and wage/salary employed are different.

## References

- Blank, Rebecca M., "Are Part-Time Jobs Bad Jobs?," in Gary Burtless, (ed.), *A Future of Lousy Jobs?*, (Washington, D.C.: Brookings Institution, 1990).
- Blau, David M., "Self-Employment, Earnings and Mobility in Peninsular Malaysia," *World Development* 14 (1986): 839-52.
- Blau, David, "Self-Employment and Self-Selection in Developing Countries," *Southern Economic Journal* 53 (1986), 351-363.
- Co, Catherine Y., Ira N. Gang, and Myeong-Su Yun, "Returns to Returning," *Journal of Population Economics*, forthcoming (1999).
- Evans, David and Linda Leighton, "Some Empirical Aspects of Entrepreneurship," in Zoltan Acs and David Audretsch (eds.), *The Economics of Small Firms: A European Challenge* (Dordrecht: Kluwer, 1990).
- Fishe, Raymond P.H., Robert P. Trost, and Philip M. Lurie, "Labor Force Earnings and College Choice of Young Women: An Examination of Selectivity Bias and Comparative Advantage," *Economics of Education Review* 1:2 (1981), 169-191.
- Frey, Maria and Janos Timar, "Hungary," in Gerhard Bosch, Peter Dawkins and Francois Michon (Eds.), *Times are Changing: Working Time in 14 Industrialised Countries* (ILO: Geneva, 1993).
- Gabor, R.I., "Small Entrepreneurship in Hungary: Ailing or Prospering?," *Acta Oeconomica* 46 (1994), 333-346.
- Gill, Andrew M., "Choice of Employment Status and the Wages of Employees and the Self-Employed: Some Further Evidence," *Journal of Applied Econometrics* 3:3 (1988), 229-34.
- Ham, John C., "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment," *Review of Economic Studies* 49:3 (1982), 335-354.
- Heckman, James, "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46:4 (1978), 931-959.
- Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica*, 47:1 (1979), 153-161.
- Iyigun, Murat and Ann Owen, "Risk, Entrepreneurship and Human Capital Accumulation," Finance and Economics Discussion Series, Federal Reserve Board, Working Paper 1997-37.
- Jovanovic, Boyan, "Selection and Evolution of Industry," *Econometrica* 50:3 (1982), 649-670.
- Kennedy, Peter E., "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 71 (1981), 802.

Kihlstrom, Richard and Jean-Jacques Laffont, "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," *Journal of Political Economy* 87 (1979) 719-748.

Koop, Gary and Dale J. Poirier, "Learning About the Across-Regime Correlation in Switching Regression Models," *Journal of Econometrics*, 78:2 (1997), 217-227.

Maddala, G. S., "Disequilibrium, Self-Selection, and Switching Models," in Zvi Griliches and Michael D. Intriligator (eds.), *Handbook of Econometrics: Volume III*, (Amsterdam: Elsevier Science Publishers B. V., 1986).

Moffitt, Robert, "An Economic Model of Welfare Stigma," *American Economic Review*, 73:5 (1983), 1023-1035.

OECD, *Social and Labour Market Policies in Hungary*, (Paris: OECD, 1995).

OECD, *OECD Economic Surveys: Hungary*, (Paris: OECD, 1991).

O'Leary, Christopher J., "Promoting Self Employment Among the Unemployed in Hungary and Poland," UpJohn Institute Staff Working Paper 99-055, 1999  
<http://www.upjohninst.org/publications/wp/9955wp.html>

Rees, Hedley and Anup Shah, "An Empirical Analysis of Self-Employment in the U.K.," *Journal of Applied Econometrics* 1:1 (1986), 95-108.

SAS Institute, *SAS/OR Technical Report: The NLP Procedure*, (Cary, NC: SAS Institute Inc., 1997).

Sik, Endre , "Measuring the Unregistered Economy in Post Communist Transition," *Eurosocietal Report* 52 (1995), Vienna, Austria.

Srivastava, V. K. and B. Bhaskara Rao, *The Econometrics of Disequilibrium Models*, (New York: Greenwood Press, 1990).

Stelcner, Morton, van der Gaag, Jacques, and Vijverberg, Wim, " A Switching Regression Model of Public-Private Sector Wage Differentials in Peru:1985-86," *Journal of Human Resources* 24:3 (1989), 545-59.

Sumner, Daniel A., "Wage Functions and Occupational Selection in a Rural Less Developed Country Setting," *Review of Economics and Statistics* 63 (1981): 513-19.

Sziracki, Gyorgy, "Hungary: Labour Market Trends and Policies," in Georg Fischer and Guy Standing (eds.), *Structural Change in Central and Eastern Europe: Labour Market and Social Policy Implications* (Paris: OECD, 1993), 95-115.

Tunali, Insan, "A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earnings Process with Remigration," *Research in Labor Economics* 8(B) (1986), 235-283.

van der Gaag and Wim Vijverberg, "A Switching Regression Model for Wage Determinants in the

Public and Private Sectors of a Developing Country,” *Review of Economics and Statistics* 70:2 (1988), 244-252.

van Ophem, Hans, “A Modified Switching Regression Model for Earnings Differentials Between the Public and Private Sectors in the Netherlands,” *Review of Economics and Statistics* 75:2 (1993), 215-224.

Vijverberg, Wim P. M., "Consistent Estimates of the Wage Equation when Individuals Choose among Income-Earning Activities," *Southern Economic Journal* 52 (1986), 1028-1042.

Vijverberg, Wim P. M., "Measuring the Unidentified Parameter of the Extended Roy Model of Selectivity," *Journal of Econometrics* 57:1-3 (1993), 69-89.

Yuengert, Andrew M., “Testing Hypotheses of Immigrant Self-Employment,” *Journal of Human Resources* 30:1 (1995), 194-204.

**Table 1. Mean Characteristics of the Sample**

<b>Both Sexes</b>	All	Labor Force	Wage Earners	Self-
Sample Size	3145	1601	1480	121
Age (Years)	40.292	37.766	37.720	38.339
Experience (Years)		19.423	19.376	19.992
Head of household (Head = 1)	0.435	0.515	0.501	0.686***
Education (Years)	10.159	11.042	11.034	11.140
Female (Women = 1)	0.524	0.497	0.509	0.339***
Marital Status (Married = 1)	0.650	0.700	0.700	0.702
Number of Children under age 6	0.248	0.235	0.239	0.198
Budapest (Living in Budapest = 1)	0.158	0.181	0.180	0.190
Monthly Earnings in Forints		17493.838	17089.318	22441.694**
Labor Force (Working = 1)	0.509			
<b>Men</b>				
	All	Labor Force	Wage Earners	Self-
Sample Size	1498	806	726	80
Age (Years)	39.505	37.836	37.660	39.438
Experience (Years)		19.747	19.609	21.000
Head of household (Head = 1)	0.696	0.804	0.795	0.888**
Education (Years)	10.201	10.948	10.935	11.063
Marital Status (Married = 1)	0.654	0.732	0.731	0.738
Number of Children under age 6	0.254	0.310	0.320	0.225
Budapest (Living in Budapest = 1)	0.156	0.167	0.164	0.200
Monthly Earnings in Forints		20016.025	19442.249	25223.038*
Labor Force (Working = 1)	0.538			
<b>Women</b>				
	All	Labor Force	Wage Earners	Self-
Sample Size	1647	795	754	41
Age (Years)	41.007	37.696	37.777	36.195
Experience (Years)		19.094	19.153	18.024
Head of household (Head = 1)	0.198	0.223	0.219	0.293
Education (Years)	10.121	11.137	11.129	11.293
Marital Status (Married = 1)	0.645	0.668	0.670	0.634
Number of Children under age 6	0.243	0.160	0.160	0.146
Budapest (Living in Budapest = 1)	0.16	0.195	0.196	0.171
Monthly Earnings in Forints		14936.753	14823.763	17014.683
Labor Force (Working = 1)	0.483			

The null hypothesis tested is that the mean of wage/salary earners is equal to that of self-employed. \*\*\*, \*\* and \* mean statistically significant at 1%, 5%, and 10%, respectively.

**Table 2. Estimates of the Selection and Switching Equations**

	Single Switching		Switching and Selection			
	Two-Step	MLE	Two-Step		MLE	
	Self-Employment	Self-Employment	Participation	Self-Employment	Participation	Self-Employment
Constant	-1.628*** (0.551)	-1.767*** (0.585)	-4.819*** (0.249)	-0.923 (1.570)	-4.841*** (0.249)	-1.402 (1.216)
Age	0.002 (0.029)	0.009 (0.030)	0.219*** (0.013)	-0.024 (0.062)	0.220*** (0.013)	-0.003 (0.050)
Age <sup>2</sup> /100	-0.004 (0.036)	-0.015 (0.037)	-0.298*** (0.016)	0.031 (0.082)	-0.299*** (0.016)	0.002 (0.066)
Head of Household	0.269** (0.124)	0.357*** (0.115)		0.271** (0.123)		0.352*** (0.116)
Education	0.009 (0.017)	0.006 (0.018)	0.131*** (0.010)	-0.004 (0.031)	0.131*** (0.011)	-0.002 (0.027)
Female	-0.192 (0.117)	-0.134 (0.111)		-0.173 (0.124)		-0.128 (0.113)
Marital Status			0.240*** (0.062)		0.226*** (0.064)	
Children < 6			0.017 (0.068)		0.005 (0.071)	
Female* Children < 6			-0.707*** (0.093)		-0.686*** (0.096)	
$\rho_{vu}$			-0.177 (0.354)		-0.107 (0.286)	

1. Standard errors are in parentheses.

2. \*\*\*, \*\* and \* mean statistically significant at 1%, 5%, and 10%, respectively.

**Table 3. Switching Regression of Log-Earnings for Wage/Salary Earners**

	Single Switching			Switching and Selection	
	OLS	Two-Step	MLE	Two-Step	MLE
Constant	8.608*** (0.055)	8.402*** (0.145)	8.579*** (0.063)	8.341*** (0.216)	8.549*** (0.132)
Budapest	0.161*** (0.027)	0.161*** (0.027)	0.159*** (0.028)	0.161*** (0.027)	0.159*** (0.028)
Education	0.080*** (0.004)	0.075*** (0.008)	0.079*** (0.005)	0.078*** (0.010)	0.080*** (0.007)
Experience	0.021*** (0.003)	0.018*** (0.006)	0.020*** (0.004)	0.020*** (0.007)	0.021*** (0.005)
Experience <sup>2</sup> /100	-0.033*** (0.008)	-0.027* (0.014)	-0.032*** (0.009)	-0.033** (0.017)	-0.035*** (0.013)
Female	-0.275*** (0.020)	-0.152** (0.073)	-0.257*** (0.021)	-0.160** (0.070)	-0.258*** (0.022)
$\lambda_1$		-1.488* (0.772)		-1.423* (0.737)	
$\lambda_2$				0.039 (0.085)	
$\sigma_e$			0.401*** (0.017)		0.401*** (0.017)
$\rho_{ev}$			-0.527*** (0.116)		-0.529*** (0.115)
$\rho_{eu}$					0.043 (0.161)
Adjusted R <sup>2</sup>	0.320	0.325		0.324	

1. Standard errors are in parentheses.

2.\*\*\*, \*\* and \* mean statistically significant at 1%, 5%, and 10%, respectively.



**Table 4. Switching Regression of Log-Earnings for Self-Employed**

	Single Switching			Switching and Selection	
	OLS	Two-Step	MLE	Two-Step	MLE
Constant	9.019*** (0.376)	10.176*** (1.422)	9.518*** (0.594)	10.065*** (1.476)	8.481*** (0.889)
Budapest	0.190 (0.153)	0.191 (0.141)	0.190 (0.142)	0.191 (0.141)	0.192 (0.137)
Education	0.053* (0.029)	0.046 (0.036)	0.050 (0.034)	0.058 (0.047)	0.093** (0.043)
Experience	0.032 (0.020)	0.029 (0.019)	0.030* (0.018)	0.038 (0.030)	0.062** (0.026)
Experience <sup>2</sup> /100	-0.083* (0.046)	-0.076 (0.046)	-0.078* (0.045)	-0.100 (0.074)	-0.163** (0.068)
Female	-0.278* (0.124)	-0.100 (0.250)	-0.203 (0.137)	-0.090 (0.251)	-0.240 (0.154)
$\lambda_1$		-0.602 (0.703)		-0.700 (0.715)	
$\lambda_2$				0.155 (0.372)	
$\sigma_e$			0.665*** (0.089)		0.763*** (0.161)
$\rho_{ev}$			-0.392 (0.265)		-0.331 (0.433)
$\rho_{eu}$					0.711*** (0.234)
Adjusted R <sup>2</sup>	0.072	0.069		0.064	
LR test for $\beta_0 = \beta_1$			5.523		6.929
LR test for $\beta_0 = \beta_1$ and $\sigma_0 = \sigma_1$			64.658		57.337

1. Standard errors are in parentheses.

2.\*\*\*, \*\* and \* mean statistically significant at 1%, 5%, and 10%, respectively.