

Discussion Paper No. 01-13

FIML Estimation of a Bivariate Probit  
Selection Rule – An Application on Firm  
Growth and Subsidisation

Frank Reize

## **Non-technical summary**

Sample selection models are widely used in microeconomic analysis. Most of these models are employed either as selection models or as dummy endogenous models to estimate the impact of e.g. an active labour market policy. So far, a mixture of these two models has rarely been used. This study aims to combine the sample selection model and the dummy endogenous treatment model using a bivariate selection rule. Moreover, estimation is carried out using full information maximum likelihood (FIML) methods to obtain fully efficient estimates. In most studies a two-step procedure is employed which yields consistent estimates, but inconsistent standard errors. The FIML estimator has rarely been used in microeconomics so far, as estimation may be complex.

To apply the FIML estimator, this study analyses the impact of the so called “bridging allowance” on firm growth. In Germany start-ups from unemployment may be supported as part of the active labour market policy with bridging allowance. Why should we apply a bivariate selection rule for this kind of investigation? First, there is the classical selection bias that only for surviving firms as well as for firms, which were investigated at least twice an employment growth rate can be derived. Both selection mechanisms are modelled within one equation. Secondly, the typical treatment sample selection bias occurs, as the subsidisation of start-ups with bridging allowance is non-random.

The empirical analysis is based on a regional sample of the ZEW Firm Start-up Panel covering 15 labour market districts using firms founded during 1993 and 1995. Estimation results show that bridging allowance reduces firm growth in Western Germany by almost 10% and in the eastern parts by 6%. These results question the efficiency of this programme. The comparison of the FIML and the two-step estimator shows that the FIML estimator is clearly warranted. The two-step procedure yields almost the same results as the FIML estimator apart from the most interesting parameter, namely bridging allowance. In the two-step procedure the impact of bridging allowance on firm growth is not significantly different from zero. This is an important result as it shows, that the consistency of the two-step estimator is not a sufficient quality for models estimating treatment effects as the significance of the effect is an essential result to assess the efficiency of a treatment. Furthermore the two-step procedure tends to neglect the correlation between the selection equations and the growth equation. Finally, the often mentioned heavy computational burdens entailed by the FIML estimator cannot be found in this study. Moreover, deriving consistent, but inefficient, standard errors for the two-step model would be even more complex.

# **FIML Estimation of a Bivariate Probit Selection Rule - An Application on Firm Growth and Subsidisation**

**Frank Reize**

*Zentrum für Europäische Wirtschaftsforschung (ZEW)  
Centre for European Economic Research*

**March 2001**

## **Abstract**

This study applies a full information maximum likelihood (FIML) estimator of the sample selection model with bivariate selection rule for the investigation of the impact of subsidised firm foundation from unemployment on employment growth of the firm. The empirical analysis is based on the ZEW Firm Start-up Panel using a cohort of firms founded in 15 labour market districts during 1993 and 1995. Estimation results show that the use of the FIML estimator is clearly warranted, compared to a two-step estimator. The FIML model yields a significant negative impact of bridging allowance on employment growth, whereas the two-step estimator underestimates the impact.

JEL-Classification: C35, D92, J68

## **Acknowledgement**

I thank François Laisney and Friedhelm Pfeiffer for valuable comments. Part of the work was financially supported by the European Social Fund.

Correspondence to: Frank Reize  
Centre for European Economic Research (ZEW)  
P.O. Box 103443  
D-68034 Mannheim  
e-mail: reize@zew.de

# 1 Introduction

There is a vast literature on sample selection models in all fields of economics. Almost all applications of these models are based on the well known Heckman selection model, either employed as a two-step estimator (Heckman 1976, 1979), or less often as full information maximum likelihood (FIML) estimator. Common to most of these estimators is that they are based on an univariate selection equation and a linear output equation. One can distinguish two major economic problems where sample selection correction is warranted. Firstly, models where output is only observed for a non-random sample of the whole population (selection model), e.g. working hours for those who are actually working. The second class of sample selection bias models arises in economic investigations of the impacts of training or other active labour market policies (treatment model). In contrast to the former models, output is not only observed for a sub-population but output (e.g. income) is influenced by the probability of belonging to a sub-population, e.g. those people receiving training would have had higher income also without training, due to specific (maybe unobserved) characteristics. Such models are also known as dummy-endogenous regression models.

In contrast to the widely used univariate selection rule, there are only a few studies employing bivariate selection (e.g. Fraker and Moffit, 1988, Goux and Maurin, 2000 and Pfeiffer or Reize, 2000a). When a bivariate selection rule is modelled, two-step estimation is used often (e.g. Goux and Maurin, 2000). However, the standard errors of the final regression are inconsistent in the two-step case and adjustment is exceedingly cumbersome. Therefore a FIML estimator, which can be applied to either pure selection models or pure treatment models or a mixture of both seems to be natural. What are the reasons for the lack of FIML models in economic literature (except Fraker and Moffit, 1988)? Two possible answers may exist (Nawata, 1994): first, maximisation of the likelihood may be difficult as the function is not globally concave, i.e. there are multiple local maxima. Second, the maximisation procedure may not converge at all if the correlation parameters are close to -1 or 1.

In this paper both estimation methods of the seldom used bivariate selection model, the two-step and the FIML estimation, are compared. Both procedures are applied on the estimation of the impact of subsidisation of firm foundation out of unemployment on employment growth of the firm.<sup>1</sup> The study intends to test, whether the computational burdens mentioned for the FIML estimator predominate the disadvantage of inconsistent standard errors of the two-step estimator, rather

---

<sup>1</sup> This paper is an extension of Pfeiffer and Reize (2000a), as it focuses on the econometric instead of the economic issue. Furthermore the database of the study is extended to a longer period of time and to larger number of firms.

than comparing the behaviour of both estimators, which should be done by a Monte Carlo simulation.

The paper is structured as follows. The next section provides the econometric model, whereas section three briefly discusses the application and data. In section four the estimation results for the FIML and the two-step model are compared, followed by the concluding section.

## 2 Econometric Modelling

### 2.1 Bivariate Selection Rule

Suppose a linear output equation for which two selection mechanisms are at work. The first selection bias results from the partial observability of the outcome variable, i.e. the outcome is only observed for a non-random sub-sample. The second selection rule arises from an endogenous dummy variable included on the right hand side of the output equation. Thus, the following simultaneous three equation model can be built up:

$$(1) \quad g_i = x_i' \mathbf{b} + d_i \mathbf{g}_g + \mathbf{e}_i,$$

$$(2) \quad d_i^* = w_i' \mathbf{d} + u_i,$$

$$(3) \quad s_i^* = z_i' \boldsymbol{\gamma} + d_i \boldsymbol{\gamma}_s + \mathbf{n}_i,$$

with

$$(4) \quad d_i = \begin{cases} 1 & \text{if } d_i^* > 0 \\ 0 & \text{otherwise} \end{cases},$$

$$(5) \quad s_i = \begin{cases} 1 & \text{if } s_i^* > 0 \\ 0 & \text{otherwise} \end{cases}.$$

The outcome variable  $g_i$  is observed iff,  $s_i=1$ . In equation (1),  $x_i$  stands for a vector of covariates with the corresponding vector of coefficients  $\mathbf{b}$  and is assumed to be independent of the error term  $\mathbf{e}_i$ .  $\mathbf{e}_i$  is assumed to be normally distributed with mean zero and variance  $\sigma_e^2$ .  $\mathbf{g}_g$  measures the impact of the endogenous dummy variable  $d_i$ . In equation (2),  $d_i^*$  is determined by a vector of characteristics  $w_i$ , whose influence is measured through the coefficient vector  $\mathbf{d}$ , while  $u_i$  is assumed to be a standard normally distributed error term, which is independent of  $w_i$ . Finally, in equation (3)  $z_i$  is a vector of variables explaining the observation mechanism for  $g_i$  and  $\mathbf{n}_i$  is assumed to be a standard normally distributed error term.  $z_i$  and  $\mathbf{n}_i$  are

assumed to be independent, whereas, like in equation (1),  $d_i$  is supposed to be endogenous. The effect of  $d_i$  is captured by  $\mathbf{g}_s$ .

The correlations between the error terms are left completely unrestricted. Therefore, the trivariate structure of the model leads to the following variance matrix:

$$(6) \quad \Sigma = \text{Var} \begin{pmatrix} \mathbf{e}_i \\ u_i \\ \mathbf{u}_i \end{pmatrix} = \begin{pmatrix} \mathbf{s}_e^2 & \mathbf{s}_{eu} & \mathbf{s}_{en} \\ \mathbf{s}_{eu} & 1 & \mathbf{s}_{in} \\ \mathbf{s}_{en} & \mathbf{s}_{in} & 1 \end{pmatrix}.$$

The identification of  $\mathbf{g}_g$  and  $\mathbf{g}_s$  is guaranteed if  $w_i$  includes at least one variable which is included neither in  $x_i$  nor in  $z_i$  and if the joint distributional assumption of the error terms is correct (see Heckman et al., 1999). Furthermore, the model is consistent iff the selection equations are recursive (see Gourieroux, 2000). This is met, as equation (2) can be expressed as a function of exogenous variables only and equation (3) as a function of exogenous variables and  $d_i^*$ .

## 2.2 The Two-step Estimator

A natural starting point for estimation would be an extension of Heckman's two-step estimator of the univariate selection model to the bivariate selection rule. In the first step, equations (2) and (3) are estimated using a bivariate probit model to obtain the inverses of Mill's ratio, which are defined as:

$$(7) \quad \mathbf{I}_{ui} = \begin{cases} \mathbf{f}[w_i' \mathbf{d}] \frac{\Phi[z_i' \mathbf{w} + d \mathbf{g}_s - \mathbf{r}_{uv} w_i' \mathbf{d}]/(1 - \mathbf{r}_{uv}^2)^{1/2}}{\Phi_2[w_i' \mathbf{d}, z_i' \mathbf{w} + d \mathbf{g}_s, \mathbf{r}_{uv}]} & \text{if } d_i = 1 \\ \mathbf{f}[-w_i' \mathbf{d}] \frac{\Phi[z_i' \mathbf{w} + d_i \mathbf{g}_s - \mathbf{r}_{uv} w_i' \mathbf{d}]/(1 - \mathbf{r}_{uv}^2)^{1/2}}{\Phi_2[-w_i' \mathbf{d}, z_i' \mathbf{w} + d \mathbf{g}_s, -\mathbf{r}_{uv}]} & \text{if } d_i = 0 \end{cases},$$

$$(8) \quad \mathbf{I}_{vi} = \begin{cases} \mathbf{f}[z_i' \mathbf{w} + d_i \mathbf{g}_s] \frac{\Phi[w_i' \mathbf{d} - \mathbf{r}_{uv} (z_i' \mathbf{w} + d_i \mathbf{g}_s)]/(1 - \mathbf{r}_{uv}^2)^{1/2}}{\Phi_2[w_i' \mathbf{d}, z_i' \mathbf{w} + d_i \mathbf{g}_s, \mathbf{r}_{uv}]} & \text{if } d_i = 1 \\ \mathbf{f}[z_i' \mathbf{w} + d_i \mathbf{g}_s] \frac{\Phi[-w_i' \mathbf{d} + \mathbf{r}_{uv} (z_i' \mathbf{w} + d_i \mathbf{g}_s)]/(1 - \mathbf{r}_{uv}^2)^{1/2}}{\Phi_2[-w_i' \mathbf{d}, z_i' \mathbf{w} + d \mathbf{g}_s, -\mathbf{r}_{uv}]} & \text{if } d_i = 0 \end{cases},$$

In the second step  $g_i$  is linearly regressed on  $x_i$ ,  $d_i$ ,  $\mathbf{I}_{ui}$  and  $\mathbf{I}_{vi}$ , which yields

$$(9) \quad g_i = x_i' \mathbf{b} + d_i \mathbf{g}_g + \mathbf{I}_{ui} \mathbf{s}_{eu} + \mathbf{I}_{vi} \mathbf{s}_{ev} + \mathbf{e}_i.$$

Ordinary least squares estimation of (9) produces consistent estimates of  $\mathbf{b}$ ,  $\mathbf{g}_g$ ,  $\mathbf{s}_{eu}$  and  $\mathbf{s}_{ev}$ , however the estimated standard errors are inconsistent. A correction of the

standard errors would be exceedingly cumbersome (see Greene, 1995). At this stage it should be noted that a zero correlation between  $u_i$  and  $n_i$  would lead to a dramatic simplification of the model (see Greene, 1998). In this case the Mill's ratios would be computed from two independent probits and correction of the standard errors would be the same as for the traditional Heckman model. But it should be also noted that solely a zero value for either  $\mathbf{s}_{eu}$  or  $\mathbf{s}_{en}$  would still lead to a complex structure of the Mill's ratio (inclusion of the bivariate normal distribution) and its derivative, respectively.<sup>2</sup> Therefore, the next step to solve our estimation problem is to proceed with FIML estimation.

### 2.3 The Full Information Maximum Likelihood Estimator

To account for the possible correlation between the three error terms, the model can be estimated in one step (i.e. fully simultaneously) using FIML techniques. In contrast to the two-step procedure, such technique will not only produce consistent, but also fully efficient estimates. As we have a bivariate selection rule, there are four different types of contribution to the log-likelihood function. Let us first define

$$(10) \quad h_d = \frac{w_i' \mathbf{d} + \frac{\mathbf{r}_{eu}}{\mathbf{s}_e} (g_i - x_i' \mathbf{b})}{\sqrt{1 - \mathbf{r}_{eu}^2}} \text{ and}$$

$$(11) \quad h_s = \frac{z_i' \mathbf{w} + \frac{\mathbf{r}_{en}}{\mathbf{s}_e} (g_i - x_i' \mathbf{b})}{\sqrt{1 - \mathbf{r}_{en}^2}},$$

where  $\mathbf{r}_{eu} = \frac{\mathbf{s}_{eu}}{\mathbf{s}_e}$  and  $\mathbf{r}_{en} = \frac{\mathbf{s}_{en}}{\mathbf{s}_e}$ . This leads to the log-likelihood function, which is defined as

---

<sup>2</sup> To correct the standard errors, the derivatives of the Mill's ratios (which are non-linear functions of the cumulative bivariate normal distribution) are needed.

$$(12) \quad \ln L = \begin{cases} \ln(\Phi_2[-w'_i \mathbf{d}, -z'_i \mathbf{w}, \mathbf{r}_{un}]) & \text{if } d_i=0 \text{ and } s_i=0 \\ \ln(\Phi_2[w'_i \mathbf{d}, -z'_i \mathbf{w}, -\mathbf{r}_{un}]) & \text{if } d_i=1 \text{ and } s_i=0 \\ \ln(\Phi_2[-\mathbf{h}_d, \mathbf{h}_s, -\mathbf{r}_{un}]) - \frac{1}{2} \ln(2\mathbf{p}) - \ln(\mathbf{s}_e) - \frac{1}{2} \left( \frac{\mathbf{g}_i - \mathbf{x}'_i \mathbf{b}}{\mathbf{s}_e} \right)^2 & \text{if } d_i=0 \text{ and } s_i=1 \\ \ln(\Phi_2[\mathbf{h}_d, \mathbf{h}_s, \mathbf{r}_{un}]) - \frac{1}{2} \ln(2\mathbf{p}) - \ln(\mathbf{s}_e) - \frac{1}{2} \left( \frac{\mathbf{g}_i - \mathbf{x}'_i \mathbf{b}}{\mathbf{s}_e} \right)^2 & \text{if } d_i=1 \text{ and } s_i=1 \end{cases} .$$

As we can see from equation (12) the extension of the univariate selection rule to a bivariate rule is straightforward.<sup>3</sup> The main differences are

- (i) the additional correlation parameter  $\mathbf{r}_{un}$  between the selection equations,
- (ii) the implementation of the bivariate normal distribution  $F_2$  instead of the normal distribution and
- (iii) two more regimes on which the output equation can be selected.

From equation (12) the mixture between the treatment and the selection model can be seen. Whereas the output regression has no contribution on the log-likelihood function if  $s_i=0$ , there is a contribution for both regimes of  $d_i$ . Finally, equation (12) shows the parameters which are estimated. Beside the coefficients, including the impact of the endogenous dummy on  $g_i$ ,  $\mathbf{g}_g$ , four more structural parameters are estimated. These are the correlation coefficients either between  $g_i$  and  $d_i^*$ ,  $\mathbf{r}_{eu}$ , or between  $g_i$  and the probability  $g_i$  is observed,  $\mathbf{r}_{en}$ . Also the correlation coefficient between the selection equations  $\mathbf{r}_{un}$  is estimated, as well as the standard error of the growth equation  $\mathbf{s}_e$ . The standard errors of the coefficients are estimated using the robust Huber/White/Sandwich estimator instead of the negative inverse of the Hessian (see e.g. Binder, 1983 or White, 1982). This estimator does not assume a correct specification of distribution of the likelihood function, but merely independence of the observations. It is defined as:

---

<sup>3</sup> For the derivation of the log-likelihood function of the traditional Heckman selection model see e.g. Amemiya (1985) or Nawata (1994). Di Tomasso (1999) employs a similar log-likelihood for a trivariate structure. However, she estimates two probits as output equations including a continuous endogenous variable.



$$(13) \quad \hat{V}(\hat{\Pi}) = H(\hat{\Pi})^{-1} \left( \frac{N}{N-1} \sum_{i=1}^N \text{grad}_i \text{grad}_i' \right) H(\hat{\Pi})^{-1},$$

where  $\hat{\Pi}$  is the vector of the estimated coefficients,  $H$  is the Hessian matrix and  $\text{grad}_i$  is the gradient vector of the  $i$ -th observation. The estimation is carried out using Stata's (Version 6.0) ML method.

### 3 Applying the Bivariate Selection Rule on Firm Growth and Subsidisation

#### 3.1 Employment Growth and Selectivity

The econometric model of the previous section is applied to estimate the impact of public subsidisation of firm foundation by the unemployed on employment growth of the start-up.<sup>4</sup> In Germany, the transition from unemployment to self-employment is subsidised as part of active labour market policies. The Federal Employment Services may pay what is known as “bridging allowance” (Überbrückungsgeld) according to §57 of the 3<sup>rd</sup> Social Security Code (SGB III).<sup>5</sup> The bridging allowance was first introduced on January 1st, 1986. Since then, it has undergone several legal changes, when periods of more liberal use were followed by periods of rather restrictive use and vice versa. August 1, 1994 represents the most recent legal change of bridging allowance, with a considerable improvement in terms of promotion. Since then, bridging allowance has generally been granted to people unemployed for at least four weeks. The duration of promotion is in general for 26 weeks and amounts for the lastly paid unemployed assistance. In addition, during the time of support the contributions to health and nursing insurance, as well as to the retirement fund, are financed at the level of the social security contributions which were last paid for the unemployed person.

To assess the quantitative impact of bridging allowance on employment growth an ordinary least squares regression may be used, including an indicator whether the firm was founded out of unemployment or not. But as known from studies investigating the direct impact of social programmes, participants and non-participants of the programme systematically differ in their observable and unobservable characteristics.<sup>6</sup> Therefore, the econometric framework of programme evaluation seems to be appropriate in this study, as the probability of starting a

---

<sup>4</sup> For a detailed discussion on the economics of firm growth and subsidisation as well as on the construction and details of the underlying database, the ZEW Firm Start-up Panel, see Pfeiffer and Reize (2000a).

<sup>5</sup> Until January 1, 1998, the legal foundation was §55a of the Labour Promotion Law (AFG).

<sup>6</sup> See e.g. Heckman et al. (1999).

business from unemployment, as well as employment growth, may be influenced by the same observable and unobservable characteristics.<sup>7</sup> Firm foundation from unemployment is allied with various selection processes, e.g. a competent authority has to assess the sustainability of the self-employment envisaged. Furthermore, there are various observable characteristics, like legal form of the firm or industry, and unobservable ones, like personal abilities and employment history of the founder, capital endowment or transition costs, which may influence both processes. As a consequence, a dummy endogenous regression model seems to be appropriate to assess the impact of bridging allowance.

The second selection problem arises from the fact that employment growth is not observed for all firms. A growth rate can only be determined for the surviving, i.e. the successful companies. However, what one tries to measure is an effect for all start-ups and not only for the group of surviving companies. An estimate using the sample of the surviving companies can overestimate effects on employment growth (“Survivor-Bias”). A further selection process might result for data reasons, since employment growth can only be calculated if the firm was examined at least twice. These two selection biases are captured by one equation.

Thus, two selection mechanisms are at work: the classical Heckman selection problem, that employment growth is only observed for a non-random sub-population and the selection bias arising from non-random firm foundation out of unemployment.

Then, the trivariate model of section (2) is employed as follows: the employment growth rate  $g$  is defined as

$$(14) \quad g = \frac{\ln E(t_2) - \ln E(t_1)}{t_2 - t_1}$$

with  $\ln E(t_t)$  being the logarithm of the number of employees at the time of examination  $t_t$ .<sup>8</sup>  $t_1$  and  $t_2$  are chosen in such a way that the point of  $t_1$  is the earliest statement on employment and  $t_2$  corresponds to the number of employees determined at the last investigation of the firm. We further assume  $d_i^*$  to be the latent index of being subsidised with bridging allowance and  $s_i^*$  to be the latent index of having a valid observation on employment growth.

---

<sup>7</sup> Firm foundation from unemployment and subsidisation with BA is synonymous, as hardly any unemployed would start a business without BA. For a detailed discussion on this, see Pfeiffer and Reize (2000a), Pfeiffer and Reize (2000b) and Reize (2000).

<sup>8</sup> This specification has been used in literature on firm growth before. See e.g. Evans (1987).

## 3.2 Data

The econometric analysis is based on a regional sample of the ZEW Firm Start-up Panel (ZFSP).<sup>9</sup> The observation unit is the legally independent enterprise and not the operational facility. The ZFSP contains newly registered enterprises in Western and Eastern Germany. The information is updated, but not in regular time intervals, and includes firm characteristics like legal form or number of employees, as well as socio-economic characteristics of shareholders, and information on a shut-down or bankruptcy.

The sample used in this study is a matched database, combining the ZFSP and a full census of bridging allowance recipients obtained by the Institute for Employment Research (IAB) in 15 labour market districts during 1994 and 1995. The sample thus includes companies started between the 4th quarter of 1993 and the first half of the 3rd quarter of 1995 in which unemployed persons receiving bridging allowance are involved, and companies for which this is not the case. Start-ups which are not eligible for subsidisation according to Labour Promotion Law (non-profit associations, agricultural enterprises) have been excluded. In addition the sample is restricted to small and medium start-ups. The aim of this step is again to exclude types of start-ups which are not applicable to the unemployed. Start-ups with 100 or more employees are usually founded by already existing firms and not by individuals.

After the selection, the sample taken from the eleven labour market districts in the old federal states includes 6,031 enterprises, of which 237 are founded by an unemployed person receiving bridging allowance. The sample from the four labour market districts in the new federal states includes 5,323 enterprises, of which 426 are promoted from unemployment.<sup>10</sup>

## 3.3 Variables and Descriptive Statistics

The factors  $w$  which characterise the heterogeneity of start-ups and which are included in the econometric specification of equation (2) are: legal form, industry, type of foundation (original or derivative), shareholder structure, unemployment to vacancy ratio and the legal arrangement of bridging allowance. These factors can be derived from the discussion on the probability of entering self-employment from employment and unemployment, respectively (see Pfeiffer and Reize, 2000a).

---

<sup>9</sup> The data in the ZFSP had been made available to the ZEW since 1989 by the Association “Verband der Vereine Creditreform” (VVC): for further details see Harhoff and Steil (1997).

<sup>10</sup> Thus we are able to identify 1,741 more firms of which 45 are subsidised compared to the previous study (Pfeiffer and Reize, 2000a). As start-ups can enter the panel with some delay, the six additional waves increase the number of firm start-ups for all cohorts.

People will enter self-employment if their expected returns from self-employment are higher than their actual returns from wage work or unemployment plus some additional transition costs. It can be shown that returns and transition costs mainly depend on human capital, financial endowment, social networks and the macro-economic situation. It seems likely that unemployed people face lower returns and higher transition costs than employed people. Therefore, firms formed by the unemployed are probably rather small, “easily” founded firms with low capital endowment and foundation is more likely during periods of high unemployment.<sup>11</sup>

The growth equation is determined by variables typically found in studies examining employment growth on the basis of the ZFSP (see e.g. Harhoff et al., 1998). Those factors  $x$  are derived either from an industrial economics approach including initial employment, legal form, industry and region, or from a labour economics approach including age and gender. Moreover, to account for organisational factors, networks and the type of foundation (original or derivative) are controlled. Finally to account for business cycles, dummy variables are included indicating the different combinations of  $t_1$  and  $t_2$  on a yearly basis. Beside the already mentioned  $x$  variables the vector  $z$  in equation (3) consists of credit worthiness and the payment history of the company. Table 1 contains a complete list of variables used for estimation as well as some descriptive statistics on these variables for start-ups with and without subsidisation through bridging allowance (BA). Due to missing values for industry, sex and age, the number of observations for estimation reduces to 4,632 for Western Germany and to 3,437 for the New Federal States.<sup>12</sup> An employment growth rate can be estimated for 2,415 start-ups in the West and 1,615 in the East.

In the Western part of Germany the average employment growth rate is 9.6% for the subsidised and 7.3% for the non subsidised start-ups (Table 1). Hence, the annual growth rate is just about a half of that in the former study on the impacts of bridging allowance (see Pfeiffer and Reize, 2000a). As the period of observation with 1,390 days for the non subsidised and 1,598 days for the subsidised firms is between two and three times higher compared to the former analysis, the average growth rate seems to turn down with growing age of the firm. From other studies (e.g. Harhoff et al., 1998 or Almus et al., 1999) it is known that young firms reach their maximum annual growth rate within the first two years after establishment. In the Eastern part, the growth rate for subsidised start-ups is 7.0% and for non-subsidised start-ups 7.8%. The period of observation is somewhat shorter than for the Western parts. Subsidised firms are observed for 1,404 days on average and non-subsidised for 1,373.

---

<sup>11</sup> See Pfeiffer and Reize (2000a) for a more detailed discussion.

<sup>12</sup> The dramatic reduction in the Eastern sample is due to over 30% missing observations for age. Estimation results excluding age do not change systematically. They are available upon request.

**Table 1: Variables and descriptive statistics**

Variables	Western Germany		Eastern Germany	
	No BA	BA	No BA	BA
All observations	4,440	192	3178	259
Non-censored observations	2,305	110	1464	151
Employment growth	0.073 (0.279)	0.096 (0.173)	0.078 (0.341)	0.070 (0.147)
Time of observation	1,390 (535)	1,540 (468)	1,373 (510)	1,404 (468)
Initial employment	2.9 (4.8)	2.0 (1.8)	3.6 (5.4)	2.4 (2.6)
Legal form				
Corporate firm	30.4	18.2	20.5	14.7
Non-corporate firm	8.1	17.2	9.6	17.4
Trade enterprise / sole proprietorship	61.4	64.6	70.0	68.0
Industry				
Construction	14.6	19.8	24.0	22.8
Manufacturing	9.8	14.6	8.1	9.3
Wholesale trade	7.7	5.7	7.2	10.0
Retail trade	24.9	26.0	19.5	24.7
Hospitality	11.2	7.3	8.8	6.2
Communication / transportation	5.1	3.7	5.5	4.3
Finance / insurance / real estate / housing	7.0	4.2	9.4	8.9
Data processing	2.9	5.7	—	—
Business related services <sup>a</sup>	7.3	8.9	8.2	8.1
Other services	9.6	4.2	9.4	5.8
Labour market district				
Other	9.9	9.9	18.2	10.8
West: Bremen; East: Schwerin	7.4	6.8	32.3	40.5
West: Hanover; East: Berlin	15.4	26.0	10.5	11.6
West: Kassel; East: Pirna	9.4	15.6	17.5	17.0
West: Essen; East: Dessau	9.6	5.2	21.5	20.1
Hof / Bayreuth	8.6	6.8	—	—
Bad Kreuznach / Mainz	17.4	5.7	—	—
Deggendorf / Landshut	11.0	10.9	—	—
Göppingen	11.4	13.0	—	—
Derivative	0.4	1.6	1.8	0.8
Diversified	13.7	19.3	16.6	16.2
Network	12.9	9.4	28.5	16.6
Female	24.1	13.0	21.7	22.4
Age				
Age < 25	7.4	8.3	10.0	6.2
25 ≤ age < 30	19.1	22.9	16.1	16.6
30 ≤ age < 35	22.2	26.0	19.0	20.1
35 ≤ age < 40	17.3	15.6	16.4	20.5
40 ≤ age < 45	13.1	10.9	16.3	15.4
45 ≤ age < 50	9.5	10.4	9.4	8.1
50 ≤ age < 55	5.9	5.2	8.5	10.0
55 ≤ age < 84	5.6	0.5	4.3	3.1
U/V-ratio	3.4 (3.5)	4.0 (3.0)	3.3 (2.8)	2.9 (1.6)
Number of additional shareholders				
No additional shareholders	73.6	68.8	79.8	71.8
One additional shareholder	20.3	22.9	15.8	22.0
Two additional shareholders	4.4	6.3	3.5	5.0
Three and more additional shareholders	1.8	2.1	0.9	1.2

**Table 1 continued**

Variables	Western Germany		Eastern Germany	
	No BA	BA	No BA	BA
Cohort of observation				
Cohort 1993 and before	4.9	1.0	1.5	0.8
Cohort 1994/1994	9.5	7.3	6.1	5.4
Cohort 1994/1995	2.7	2.6	1.8	1.5
Cohort 1994/1996	3.0	3.7	2.6	1.5
Cohort 1994/1997	3.1	2.6	3.2	2.7
Cohort 1994/1998	3.6	5.2	4.2	4.6
Cohort 1994/1999	8.6	7.3	7.8	8.9
Cohort 1994/2000	5.5	8.3	5.1	5.8
Cohort 1995/1995	10.8	12.5	7.7	10.4
Cohort 1995/1996	4.0	3.7	3.1	2.7
Cohort 1995/1997	4.6	3.7	3.9	5.0
Cohort 1995/1998	5.1	3.7	6.0	8.5
Cohort 1995/1999	11.5	19.3	11.4	13.9
Cohort 1995/2000	7.5	10.9	8.0	9.3
Cohort 1996/1996	2.1	1.6	5.9	2.7
Cohort 1996/1997	0.6	0.5	1.8	1.2
Cohort 1996/1998	1.2	1.0	2.7	1.5
Cohort 1996/1999	3.2	2.1	6.2	5.8
Cohort 1996/2000	1.7	1.0	3.3	3.5
Cohort 1997/1997	1.3	0.5	2.2	1.2
Cohort 1997/1998	0.4	0.0	1.0	0.0
Cohort 1997/1999	1.1	0.5	1.8	1.5
Cohort 1997/2000	0.5	0.0	1.3	0.8
Cohort 1998 and later	3.6	1.0	1.5	0.8
Quarter of foundation (Legal Arrangement of BA)				
Quarter 93/4	12.9	4.7	11.6	7.7
Quarter 94/1	10.9	12.5	15.6	13.1
Quarter 94/2	17.0	13.5	16.2	9.3
Quarter 94/3	7.6	6.8	11.9	13.1
Quarter 94/4	13.9	23.4	10.9	22.8
Quarter 95/1	14.9	27.6	13.5	23.9
Quarter 95/2 <sup>b</sup>	12.2	9.9	20.4	10.0
Quarter 95/3	10.7	1.6	—	—
Payment histories				
Pays bills without delay <sup>c</sup>	47.8	55.2	46.4	59.5
No experience with respect to payment behaviour	18.6	17.2	13.4	17.4
Payment within 30 days	2.3	1.0	—	—
Payment takes longer	3.2	2.6	3.9	0.4
Pays slowly	9.8	10.4	9.1	6.2
Payment after reminder	2.0	2.1	3.5	1.9
Payments overdue	16.4	11.5	23.0	14.3
Credit worthiness				
Credit possible or advised	42.8	50.0	16.7	18.2
No credit experience	22.7	21.4	19.7	21.2
Limited credit	14.0	13.5	35.1	46.3
Secured credit	3.0	1.0	3.6	0.8
No credit recommended	17.4	14.1	25.0	13.5

Notes: Regional sample of the ZFSP; mean for continuous and share in %-points for discrete variables; standard deviation for continuous variables in parenthesis; for Eastern Germany <sup>a</sup> data processing and business related services, <sup>b</sup> quarter 95/2 and quarter 95/3 as well as <sup>c</sup> pays bills without delay and payment within 30 days are combined.

## 4 Estimation Results

The estimation results for Western Germany show only marginal differences between the FIML estimator and the two-step procedure (Table 2). It is not surprising that the coefficients are nearly identical, but also the standard errors in the two-step model, which are not corrected, have about the same values as the consistent and efficient errors obtained by FIML estimation.<sup>13</sup> As a consequence, should we forget about FIML estimation and just proceed with the less complex two-step procedure without correction of the standard errors? The answer is no. Upon a closer examination of the results we find two major differences between the two models. First, the most interesting coefficient in the growth equation, namely bridging allowance, is negative and significant in the FIML model and negative but insignificant in the two-step model. Furthermore the coefficient in the two-step model is around 10% smaller than the coefficient obtained by FIML estimation. The second major difference between the two estimators, which is closely related to the first one, is the insignificant correlation between employment growth and subsidisation in the two-step model compared to a significant and positive correlation in the FIML model.<sup>14</sup> This result suggests a positive influence of unobservable characteristics on the probability of starting a firm with the support of bridging allowance as well as on employment growth. The two-step model underestimates this unobserved impact and therefore leads to an upward biased (from negative towards zero) estimate of bridging allowance. Though the estimates are still consistent in the two-step model, the inefficiency of this procedure may lead to wrong conclusions. This seems to be a serious problem especially for dummy endogenous models, as the most interesting parameter of these models, the endogenous dummy, is most affected by the inefficiency. In our case the sole use of the two-step estimator would have led to the result that bridging allowance has no impact on firm growth, whereas bridging allowance in fact reduces growth by almost 10%.

Similar results are obtained for Eastern Germany. Again almost all coefficients and standard errors are the same for both estimators, apart from the coefficient on bridging allowance and the correlation between subsidisation and growth. As for Western Germany, both estimates in the two-step model are insignificant and

---

<sup>13</sup> It should be well noted, that the standard errors of the two-step estimation are only inconsistent for the growth equation. Those in the first step (the bivariate selection equations) are also obtained using FIML estimation. Therefore, the standard errors of the selection equations can even be smaller for the two-step estimation compared to the FIML estimation.

<sup>14</sup> Table 2 and Table 3 show the estimated coefficients of the inverses of Mill's ratio ( $\mathbf{s}_{eu}$  and  $\mathbf{s}_{ev}$ ) for the two-step model, instead of  $\mathbf{r}_{eu}$  and  $\mathbf{r}_{ev}$ . The correlation coefficients are defined as

$$\mathbf{r}_{eu} = \frac{\mathbf{s}_{eu}}{\mathbf{s}_e} \text{ and } \mathbf{r}_{ev} = \frac{\mathbf{s}_{ev}}{\mathbf{s}_e} \text{ (see section 2).}$$

slightly smaller than in the FIML model. In the eastern parts of Germany bridging allowance reduces employment growth by nearly 6%, which is about four percentage points lower than in the West.

There is a negative correlation between growth and selection in Western Germany as well as in Eastern Germany. However, the coefficient for the East is not significantly different from zero. This result may be a hint that worse performing firms are more likely examined by the VVC and that this effect predominates the positive “survivor bias”, at least in Western Germany. The correlation between selection and subsidisation is also negative, but only significant for Western Germany. Again there can be a lot of speculation on unobservable factors, which may influence selection and subsidisation in opposite directions. More important for estimation purposes is the fact that the model for the New Federal States can be reduced to a univariate treatment model, as both the correlation between subsidisation and selection as well as the correlation between selection and growth are not statistically different from zero. But, it should be well noted that the two-step estimation would still lead to the wrong conclusion that employment growth can be estimated by ordinary least squares. Again, this would yield an (overestimated) non-zero impact of bridging allowance on employment growth.

The remaining results can be briefly summarised. In both parts of Germany we find evidence against Gibrat’s Law that firm growth is proportional to firm size. Furthermore, non-corporate firms seem to grow fastest and start-ups in the finance/ insurance / real estate / housing sector slowest. Organisational factors have no effect on firm growth in the Old Federal States and in the New States start-ups within networks seem to grow faster by 5%. Finally personal characteristics influence firm growth neither in the West nor in the East.

Turning to the subsidisation equation we find for Western Germany, as expected, a positive impact of a tight labour market on the probability of firm foundation out of unemployment. Also, as expected, subsidised start-ups are more often founded as non-corporate firms in the construction sector. An interesting result is that subsidised firms have more shareholders compared to the non-subsidised. This may be a hint, that the subsidisation of an unemployed shareholder can be also attractive for formerly employed firm founders. The inclusion of an unemployed in firm foundation may increase the capital endowment of the firm. Finally during the fourth quarter of 1994 the probability of subsidisation is the highest. For Eastern Germany similar results can be found, apart that the U/V-ratio and industry have no influence on subsidisation.

As the coefficients of the selection equation are difficult to interpret because of two selection processes of firm survival and firm examination (see section 3), they will not be discussed.



**Table 2: Estimation Results for Western Germany**

Variable	FIML estimation		two-step estimation	
	Coefficient	Standard error	Coefficient	Standard error
<b>Growth equation</b>				
Bridging allowance	-0.098	0.031	-0.089	0.084
Initial firm size	-0.174	0.028	-0.176	0.028
Initial firm size squared	0.019	0.007	0.019	0.007
Non-corporate firm	-0.067	0.015	-0.072	0.015
Trade enterprise / sole proprietorship	-0.045	0.012	-0.046	0.012
Manufacturing	-0.017	0.019	-0.019	0.019
Wholesale trade	-0.025	0.017	-0.023	0.017
Retail trade	-0.055	0.014	-0.054	0.014
Hospitality	-0.029	0.019	-0.024	0.019
Communication / transportation	0.000	0.020	0.003	0.020
Finance / insurance / real estate / housing	-0.086	0.024	-0.083	0.024
Data processing	-0.033	0.024	-0.037	0.024
Business related services	-0.045	0.021	-0.044	0.021
Other services	-0.047	0.026	-0.042	0.027
Derivative	0.004	0.031	-0.003	0.032
Diversified	-0.005	0.012	-0.005	0.012
Networks	0.027	0.019	0.027	0.020
Female	-0.017	0.014	-0.016	0.014
25 ≤ age < 30	0.040	0.036	0.039	0.036
30 ≤ age < 35	0.037	0.036	0.037	0.037
35 ≤ age < 40	0.022	0.037	0.021	0.037
40 ≤ age < 45	0.019	0.038	0.019	0.039
45 ≤ age < 50	0.016	0.039	0.016	0.040
50 ≤ age < 55	0.019	0.041	0.019	0.042
55 ≤ age < 84	0.043	0.064	0.044	0.065
Constant	0.232	0.054	0.235	0.056
<b>Subsidisation equation</b>				
U/V-ratio	0.021	0.008	0.021	0.008
Non-corporate firm	0.546	0.130	0.551	0.131
Trade enterprise / sole proprietorship	0.515	0.153	0.503	0.151
Manufacturing	0.050	0.122	0.055	0.123
Wholesale trade	-0.267	0.153	-0.267	0.154
Retail trade	-0.108	0.103	-0.108	0.103
Hospitality	-0.367	0.142	-0.367	0.142
Communication / transportation	-0.264	0.180	-0.269	0.181
Finance / insurance / real estate / housing	-0.254	0.175	-0.248	0.177
Data processing	0.261	0.176	0.256	0.177
Business related services	-0.003	0.143	-0.003	0.143
Other services	-0.416	0.164	-0.408	0.161
Derivative	0.602	0.338	0.598	0.339
One additional shareholder	0.247	0.171	0.228	0.170
Two additional shareholders	0.478	0.203	0.465	0.202
Three and more additional shareholders	0.668	0.271	0.634	0.268
Quarter 93/4	-0.815	0.159	-0.794	0.155
Quarter 94/1	-0.253	0.120	-0.262	0.121
Quarter 94/2	-0.396	0.115	-0.397	0.116
Quarter 94/3	-0.415	0.142	-0.417	0.142
Quarter 95/1	-0.036	0.101	-0.030	0.101
Quarter 95/2	-0.454	0.121	-0.450	0.122
Quarter 95/3	-1.079	0.221	-1.087	0.220
Constant	-1.856	0.153	-1.846	0.152

**Table 2 continued**

Variable	FIML estimation		Two-step estimation	
	Coefficient	Standard error	Coefficient	Standard error
<b>Selection equation</b>				
Bridging allowance	1.393	0.197	1.363	0.218
Initial firm size	0.249	0.108	0.264	0.109
Initial firm size squared	-0.012	0.022	-0.015	0.021
Non-corporate firm	-0.226	0.097	-0.226	0.097
Trade enterprise / sole proprietorship	-0.169	0.063	-0.168	0.063
Manufacturing	0.175	0.099	0.176	0.100
Wholesale trade	0.175	0.106	0.174	0.106
Retail trade	0.025	0.079	0.023	0.079
Hospitality	-0.166	0.098	-0.169	0.098
Communication / transportation	0.064	0.129	0.065	0.130
Finance / insurance / real estate / housing	-0.042	0.113	-0.041	0.114
Data processing	0.165	0.144	0.170	0.145
Business related services	-0.079	0.111	-0.080	0.112
Other services	-0.346	0.105	-0.350	0.106
Derivative	-0.473	0.299	-0.469	0.299
Diversified	0.062	0.069	0.061	0.069
Networks	-0.135	0.076	-0.137	0.076
Female	-0.104	0.057	-0.104	0.057
25 ≤ age < 30	0.162	0.096	0.168	0.096
30 ≤ age < 35	0.174	0.093	0.178	0.093
35 ≤ age < 40	0.136	0.099	0.139	0.099
40 ≤ age < 45	0.074	0.105	0.079	0.105
45 ≤ age < 50	0.059	0.110	0.063	0.110
50 ≤ age < 55	0.013	0.127	0.018	0.127
55 ≤ age < 84	-0.246	0.124	-0.240	0.124
No experience with respect to payment behaviour	-0.927	0.093	-0.934	0.093
Payment within 30 days	0.504	0.206	0.494	0.208
Payment takes longer	-0.242	0.162	-0.244	0.162
Pays slowly	-0.111	0.078	-0.109	0.078
Payment after reminder	-0.379	0.253	-0.383	0.253
Payments overdue	-1.749	0.314	-1.756	0.313
No credit experience	-1.733	0.086	-1.726	0.087
Limited credit	-0.174	0.071	-0.170	0.071
Secured credit	-1.003	0.175	-1.000	0.175
No credit recommended	-1.163	0.295	-1.161	0.295
Constant	0.945	0.147	0.937	0.147
$\rho_{\text{eu}} / \sigma_{\text{eu}}$	0.194	0.037	0.050	0.040
$\rho_{\text{ev}} / \sigma_{\text{ev}}$	-0.103	0.043	-0.032	0.016
$\rho_{\text{uv}}$	-0.683	0.096	-0.669	0.109
$\sigma_{\text{e}}$		0.266		0.268
Log-likelihood		-2,541.792		—
Wald test / F test		$\chi^2(55) = 233.46$		$F(57, 2,357) = 3.98$
R-squared		—		0.077
Full sample observations			4,632	
Non-censored observations			2,416	

Notes: (i) Regional sample of the ZFSP. (ii) Reference categories for dummy variables are: corporate firm, construction, age < 25, quarter 94/4, no additional shareholders, payment without delay, credit advised or possible. (iii) Estimation includes 23 dummies for different combinations of t1 and t2 in the growth equation and 8 regional dummies in the growth and selection equation. (iv) Robust standard errors using Huber/White/Sandwich estimator of variance covariance matrix. (v) Correlation coefficient for FIML estimation and covariance coefficient for two-step estimation.

**Table 3: Estimation results for Eastern Germany**

Variable	FIML estimation		Two-step estimation	
	Coefficient	Standard error	Coefficient	Standard error
<b>Growth equation</b>				
Bridging allowance	-0.059	0.030	-0.006	0.063
Initial firm size	-0.276	0.028	-0.277	0.029
Initial firm size squared	0.049	0.007	0.049	0.007
Non-corporate firm	-0.049	0.020	-0.052	0.020
Trade enterprise / sole proprietorship	-0.044	0.020	-0.045	0.020
Manufacturing	0.041	0.035	0.040	0.035
Wholesale trade	-0.032	0.027	-0.033	0.028
Retail trade	-0.055	0.017	-0.057	0.017
Hospitality	0.033	0.061	0.034	0.062
Communication / transportation	0.076	0.057	0.076	0.058
Finance / insurance / real estate / housing	-0.123	0.021	-0.123	0.021
Business related services incl. data processing	-0.085	0.034	-0.085	0.035
Other services	-0.064	0.022	-0.063	0.022
Derivative	-0.001	0.035	0.001	0.036
Diversified	0.004	0.018	0.003	0.019
Networks	0.050	0.024	0.050	0.024
Female	-0.021	0.016	-0.021	0.016
25 ≤ age < 30	-0.013	0.029	-0.013	0.029
30 ≤ age < 35	-0.039	0.025	-0.039	0.025
35 ≤ age < 40	-0.048	0.027	-0.048	0.027
40 ≤ age < 45	-0.056	0.023	-0.056	0.023
45 ≤ age < 50	-0.050	0.027	-0.050	0.027
50 ≤ age < 55	-0.012	0.040	-0.012	0.041
55 ≤ age < 84	-0.068	0.043	-0.068	0.044
Constant	0.225	0.043	0.224	0.045
<b>Subsidisation equation</b>				
U/V-ratio	-0.014	0.012	-0.015	0.012
Non-corporate firm	0.308	0.133	0.303	0.133
Trade enterprise / sole proprietorship	0.473	0.152	0.472	0.153
Manufacturing	0.120	0.132	0.122	0.132
Wholesale trade	0.243	0.132	0.244	0.132
Retail trade	0.157	0.099	0.159	0.099
Hospitality	-0.175	0.145	-0.173	0.144
Communication / transportation	-0.104	0.166	-0.102	0.165
Finance / insurance / real estate / housing	0.043	0.128	0.043	0.128
Business related services incl. data processing	0.040	0.136	0.044	0.137
Other services	-0.179	0.142	-0.178	0.142
Derivative	-0.474	0.338	-0.475	0.338
One additional shareholder	0.490	0.180	0.494	0.181
Two additional shareholders	0.613	0.226	0.617	0.227
Three and more additional shareholders	0.575	0.348	0.571	0.347
Quarter 93/4	-0.572	0.143	-0.574	0.144
Quarter 94/1	-0.432	0.128	-0.437	0.129
Quarter 94/2	-0.624	0.140	-0.624	0.142
Quarter 94/3	-0.355	0.123	-0.355	0.123
Quarter 95/1	-0.075	0.126	-0.078	0.129
Quarter 95/2-3	-0.723	0.126	-0.723	0.126
Constant	-1.538	0.185	-1.535	0.187

**Table 3 continued**

Variable	FIML estimation		Two-step estimation	
	Coefficient	Standard error	Coefficient	Standard error
<b>Selection equation</b>				
Bridging allowance	0.926	0.613	0.898	0.667
Initial firm size	0.373	0.116	0.374	0.116
Initial firm size squared	-0.050	0.025	-0.050	0.025
Non-corporate firm	-0.272	0.112	-0.272	0.113
Trade enterprise / sole proprietorship	-0.345	0.088	-0.346	0.088
Manufacturing	-0.108	0.110	-0.107	0.110
Wholesale trade	-0.615	0.122	-0.615	0.122
Retail trade	-0.300	0.091	-0.300	0.091
Hospitality	-0.653	0.124	-0.655	0.124
Communication / transportation	-0.344	0.132	-0.345	0.133
Finance / insurance / real estate / housing	-0.384	0.113	-0.385	0.113
Business related services incl. data processing	-0.352	0.114	-0.353	0.114
Other services	-0.304	0.116	-0.305	0.117
Derivative	0.180	0.226	0.178	0.227
Diversified	0.081	0.078	0.081	0.078
Networks	0.030	0.071	0.030	0.071
Female	-0.150	0.071	-0.151	0.071
25 ≤ age < 30	0.090	0.113	0.090	0.113
30 ≤ age < 35	0.128	0.109	0.129	0.110
35 ≤ age < 40	0.132	0.109	0.132	0.109
40 ≤ age < 45	0.219	0.113	0.219	0.113
45 ≤ age < 50	0.111	0.125	0.111	0.125
50 ≤ age < 55	0.055	0.129	0.054	0.129
55 ≤ age < 84	0.021	0.174	0.021	0.174
No experience with respect to payment behaviour	-0.979	0.112	-0.979	0.112
Payment takes longer	-0.207	0.160	-0.207	0.161
Pays slowly	0.037	0.090	0.037	0.090
Payment after reminder	-0.503	0.227	-0.506	0.227
Payments overdue	-2.229	0.246	-2.236	0.247
No credit experience	-1.562	0.117	-1.565	0.117
Limited credit	-0.128	0.084	-0.128	0.084
Secured credit	-0.694	0.173	-0.696	0.174
No credit recommended	-0.914	0.237	-0.913	0.238
Constant	1.124	0.184	1.128	0.186
$\rho_{\text{eu}} / \sigma_{\text{eu}}$	0.083	0.036	-0.002	0.035
$\rho_{\text{ev}} / \sigma_{\text{ev}}$	-0.025	0.044	-0.006	0.020
$\rho_{\text{uv}}$	-0.342	0.336	-0.328	0.364
$\sigma_{\text{e}}$	0.298		0.303	
Log-likelihood	-2,471.991		—	
Wald test / F test	$\chi^2(51) = 218.01$		F(53, 1561) = 4.02	
R-squared	—		0.170	
Full sample observations	—		3,437	
Non-censored observations	—		1,615	

Notes: (i) Regional sample of the ZFSP. (ii) Reference categories for dummy variables are: corporate firm, construction, age < 25, quarter 94/4, no additional shareholders, payment without delay or within 30 days, credit advised or possible. (iii) Estimation includes 23 dummies for different combinations of t1 and t2 in the growth equation and 4 regional dummies in the growth and selection equation. (iv) Robust standard errors using Huber/White/Sandwich estimator of variance covariance matrix. (v) Correlation coefficient for FIML estimation and covariance coefficient for two-step estimation.

## 5 Conclusions

The paper presented an application of a FIML estimator for a mixture of a selection and treatment model with a bivariate selection rule. Furthermore, this estimator was compared to the classical two-step procedure. Estimation of the impact of firm subsidisation with bridging allowance on employment growth showed that the FIML estimator is clearly warranted. The FIML estimator yields a negative and significant impact of bridging allowance on firm growth, whereas the two-step procedure yields a slightly smaller and not significant effect for both parts of Germany. Though, the two-step procedure estimates coefficients consistently, it could still be misleading in models assessing the impact of active labour market policies as it produces inconsistent estimates of the standard errors. In this kind of studies not only the consistency of an estimator is important, but also the efficiency, as the researcher or the political decision-makers are interested in whether the impact is significant different from zero or not.

The two-step estimator is often proposed because of its easier usage compared to the FIML estimator. For sure, the likelihood-function in a selection model with bivariate selection rule may be complex and computation can be difficult some times, but within modern econometric software packages these burdens should be much lower (at least in this study) than the burdens of deriving the correct standard errors in the two-step model. Moreover, these corrected standard errors would still be larger than those of the FIML estimator. Hence, there seems to be no reason to use the two-step procedure instead of the FIML estimator. Though, it should be noted that there may exist other estimators (e.g. GMM), which demand less restrictive assumptions than FIML. The performance of these estimators can be compared in a Monte Carlo study, but this is beyond the scope of this paper.

A further conclusion can be drawn on the active labour market programme of bridging allowance. For Western Germany subsidised start-ups from unemployment have a 10% lower growth rate than other start-ups. For the eastern parts the growth rate is 6% smaller. This questions the efficiency of the programme and confirms our previous findings (see Pfeiffer and Reize, 2000a,b and Reize, 2000). Bridging allowance seems to fail (at least) the aim of creating additional employment through firm foundation, which could be a fact of programme design. As discussed in Pfeiffer and Reize and in Reize opportunistic or myopic behaviour of bridging allowance recipients cannot be excluded. To answer the question, whether bridging allowance reaches its direct aim, the creation of stable employment for the former unemployed, firm survival has to be analysed. But as Pfeiffer and Reize (2000a) show, this aim seems to be failed, too. Bridging allowance is a non-refundable grant, which in fact prolongs the entitlement period of unemployment benefits by six months. Therefore, businesses may be built up to receive additional unemployment benefits instead of starting a growing company. As a consequence, policy makers may think of revising bridging allowance to a (partly) refundable loan.

## **Bibliography**

- Almus, M., E. Nerlinger and F. Steil, 1999, Growth Determinants of Start-ups in Eastern Germany. A Comparison between Innovative and Non-Innovative Firms, in: Oakey, R. and W. Durning (eds.), *New Technology-Based Firms in the 1990s*, 6, London.
- Amemiya, T., 1985, *Advanced Econometrics*, Harvard University Press, Cambridge.
- Binder, D. A., 1983, On the variances of asymptotically normal estimators from complex surveys, *International Statistical Review*, 51, 279-292.
- Di Tommaso, M. L., 1999, A trivariate model of participation, fertility and wages: the Italian case, *Cambridge Journal of Economics*, 23, 623-640.
- Evans, D., 1987, Tests of Alternative Theories of Firm Growth, *Journal of Political Economy*, 95, 4, 657-674.
- Fraker, T and R. Moffit, 1988, The effect of food stamps on labor supply: a bivariate selection model, *Journal of Public Economics*, 35, 1, 25-56.
- Gourieroux, C., 2000, *Econometrics of Qualitative Dependent Variables*, University Press, Cambridge.
- Goux, D and E. Maurin, 2000, Returns to firm-provided training: evidence from French worker-firm matched data, *Labour Economics*, 7, 1, 1-19.
- Greene, W., 1995, *Limdep Version 7.0: Users Manual*, Econometric Software, Bellport, New York.
- Greene, W., 1998, Gender Economics Courses in Liberal Arts Colleges: Further Results, *Journal of Economic Education*, 29, 4, 291-300.
- Harhoff, D. and F. Steil, 1997, Die: ZEW-Gründungspanels – Konzeptionelle Überlegungen und Analysepotential, in: D. Harhoff, ed., *Unternehmensgründungen - Empirische Analysen für die alten und neuen Bundesländer*, Nomos, Baden-Baden 11 - 29.
- Harhoff, D., K. Stahl and M. Woywode, 1998, Legal Form, Growth and Exit of West German Firms - Empirical Results for Manufacturing, Construction, Trade and Service Industries, *The Journal of Industrial Economics*, XLVI, 4, 453 - 488.
- Heckman, J., 1976, The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement*, 5, 475-492.
- Heckman, J., 1979, Sample Selection Bias as a Specification Error, *Econometrica*, 47, 153-161.

- Heckman, J. J., R. LaLonde, and J. A. Smith, 1999, The Economics and Econometrics of Active Labor Market Programs, O. Ashenfelter and D. Card, eds., The Handbook of Labor Economics, Volume III.
- Nawata, K., 1994, Estimation of sample selection bias models by the maximum likelihood estimator and Heckman's two-step estimator, *Economics Letters*, 45, 33-40.
- Pfeiffer, F. and F. Reize, 2000a, From Unemployment to Self-employment – Public Promotion and Selection, *International Journal of Sociology*, 30, 3, 71-99.
- Pfeiffer, F. and F. Reize, 2000b, Business Start-ups by the Unemployed - an Econometric Analysis Based on Firm Data, *Labour Economics*, 7, 5, 629-663.
- Reize, F. ,2000, Leaving Unemployment for Self-employment. A Discrete Duration Analysis of Determinants and Stability of Self-employment among Former Unemployed, ZEW Discussion Paper, 00-26, ZEW, Mannheim.
- White, H., 1982, Maximum likelihood estimation of misspecified models, *Econometrica*, 50, 1-25.