

# Evaluating the Impact of Public Start-up Assistance – Results from an Econometric Approach <sup>†</sup>

by

MATTHIAS ALMUS

Centre for European Economic Research (ZEW), Mannheim

**Abstract:** This paper analyzes the medium-term growth performance of firms that exclusively received start-up assistance from programs administered by the Deutsche Ausgleichsbank (DtA), a state owned bank, within two years after start-up. I apply a parametric selection approach that controls for two potential sources of self selection (receipt of assistance, complete interview data). The empirical analysis shows that firms receiving assistance perform better in terms of employment growth over a six year period.

**Keywords:** New Firms, Employment Growth, Evaluation, Selection Bias, Bivariate and Simultaneous Probit.

**JEL Classification:** C35, D92, L21.

Word Count: 7,566

address: ZEW, Centre for European Economic Research,  
Department of Industrial Economics and International Management  
P.O.Box 10 34 43, D-68034 Mannheim  
phone: +49/621-1235-185  
fax: +49/621-1235-170  
E-mail: [almus@zew.de](mailto:almus@zew.de)

---

<sup>†</sup>Financial support from the Deutsche Ausgleichsbank (DtA) and the German Science Foundation under the grant LE1147/1-1 is gratefully acknowledged. Thanks to Wolfgang Franz, Georg Licht, Susanne Prantl, Frank Reize, Daniel Skambracks and Jochen Struck for valuable comments and discussion. Moreover, I thank Ruth Dollard for proof-reading. All errors that remain are, of course, the responsibility of the author alone.

**Non-technical summary:**

The Deutsche Ausgleichsbank (DtA), a state owned bank, provides with its programs in the 'Entrepreneurship/Start-up' segment more than 80 per cent of all in Germany allocated public assistance means in this segment in the past decade. An evaluation of whether these programs have an impact on the medium-term development of firms or not is necessary and wanted because of the huge amount of means spent and the high expectations connected with these programs.

This study uses information of the ZEW Entrepreneurship Study and a DtA data set that contains all assistance acceptances from 1990 to 1999 to evaluate the impact of the DtA assistance segment 'Entrepreneurship/ Start-up'. This was already done in some previous studies. Contrary to the short term horizon of these studies this analysis has a medium term evaluation horizon since employment growth over a six year period serves as success measure. Moreover, this study avoids some methodological problems of previous studies. An approach that controls for two possible sources of selection, which could lead to biased results if not considered, is therefore applied to assess whether DtA assistance has a medium term influence on the employment development. One possible source of selection concerns the fact that not the same amount of information for all firms is available in the ZEW Entrepreneurship Study. Therefore, it is possible that certain firms are more likely to possess a higher information content that can be used in the analysis. On the other hand, a descriptive analysis showed that the allocation of the DtA funds does not follow a principle of 'equal shares for all', but special firm groups are more likely to receive funding. This separates the observations into four groups. The econometric analysis shows that assistance receiving firms grow, other things equal, on average seven percentage points faster compared to firms not receiving DtA assistance. This result indicates the success of the programs under consideration.

# 1 Introduction

Evaluation of programs using econometric and statistical methods is nowadays widely accepted to assess the efficiency of means spent on public programs or measures (Heckman and Hotz 1989). This mainly concerns the evaluation of labor market programs, e.g. the evaluation of active labor market programs (ALMP) or qualification measures (LaLonde 1986, Heckman et al. 1997, Heckman and Smith 1997, Dehejia and Wahba 1999, Lechner 1999). In these cases, the subject of the examination is people and interesting research questions include whether wages, salaries or the probability of being hired or re-employed increases if the person takes part in a specific measure or program.

In contrast, the evaluation of government programs addressing firms or businesses by econometric or statistical methods has not attracted that much attention. Here, the firm or business is the object of the examination. Existing evaluation studies deal, for example, with the effect of public R&D subsidies. Klette et al. (2000) surveyed the literature dealing with public R&D subsidies and its effects on private innovation activities. David et al. (2000) review the literature on the relation of R&D subsidies and R&D expenditures. Furthermore, several studies examine the effects of SBIR<sup>1</sup> grants in the U.S. on various measures of success (Lerner 1999, Wallsten 2000). Lerner, for example, shows that the SBIR awardees achieve substantially higher employment and sales growth compared to the set of matched firms indicating the success of the SBIR program.

But there are only few studies that deal with potential differences in the probability of survival and employment growth prospects, given that new firms received public start-up assistance. However, this is an interesting question and one which should be investigated. On the one hand, huge amounts of mo-

---

<sup>1</sup> More information regarding the Small Business Innovation Research (SBIR) program provides Lerner (1999).

ney are spent every year on public support schemes to promote new firms. On the other hand, there are several expectations that are placed on new firms. First of all, start-ups are expected to create new jobs at least for the entrepreneur or the firm owner(s) themselves. This should help to overcome current labor market problems. Moreover, new firms often start with new or improved products and services. Therefore, these firms should push ahead technological change and support the transformation from an industry based economy to a service and information based one. Finally, start-ups lead to a rejuvenation of the economy. However, the generous support new firms with public assistance is justified only if and only if new firms can fulfill these expectations.

Studies in this field include Pfeiffer and Reize (2000) or Battistin et al. (2001). Battistin et al. evaluate the effect of promotion for young Italian entrepreneurs on the survival chances of their businesses. The main result of this study is that promoted firms do not have significantly better survival chances. Pfeiffer and Reize use an econometric approach to assess whether start-ups founded by unemployed people in Eastern and Western Germany that received bridging allowances (“Überbrückungsgeld”) have better survival and growth chances compared to unemployed entrepreneurs which did not receive this kind of support. While start-ups in Eastern Germany whose unemployed owners received bridging allowances have a lower survival probability, there is no effect observable in Western Germany. Moreover, bridging allowances do not influence the employment growth neither in Eastern nor in Western Germany.

This study evaluates the effect of public start-up assistance administered by the Deutsche Ausgleichsbank (DtA) on the subsequent employment growth of Eastern and Western German start-ups between 1990 and 1993. The

DtA administered the main share of German public start-up assistance<sup>2</sup> and supports entrepreneurs and business/firm start-ups in the segment ‘Entrepreneurship/ Start-up’ to maintain or even raise the number of new jobs created by new firms by increasing their survival probability as well as to help overcome capital market imperfections. The ERP<sup>3</sup> equity capital assistance program as well as the ERP and the DtA business start-up program are the most important schemes which provide firms with additional investment means.

A significant source of the high failure probability of start-ups and the resulting missing potential to create more employment are several imperfections on the German capital market. These imperfections seriously impede the start-up of new firms even in the case of a promising business concept. On the one hand, potential entrepreneurs often cannot raise the optimal amount of equity and outside capital to carry out the necessary investments. On the other, some entrepreneurs have to pay high interest rates and risk premiums on the borrowed capital where the amount depends on characteristics of the potential entrepreneurs and the firms to be funded. Harhoff and Koerting (1998) for example show that small or Eastern German firms are confronted with higher capital market costs compared to Western German or large firms.

This study aims to evaluate whether financial start-up assistance by the DtA significantly increases the medium-term employment growth rates of assisted firms. Previous studies related to this topic found ambiguous effects (see for example Brüderl et al. 1993). However, these studies exhibit several methodological problems. Firstly, the data used only allows evaluations of the short term impact and therefore these studies cannot clearly separate bet-

---

<sup>2</sup> According to DtA statements, the bank possesses a market share of 89 and 88 per cent in the years 1997 and 1998 in the segment ‘Entrepreneurship/Start-up’ (Existenz-/Unternehmensgründung) (excl. KfW [Kreditanstalt für Wiederaufbau, a state owned German bank] start-up assistance) based on the whole amount of loans in this segment. These figures are not available for the beginning of the 1990s, but may be similar or even higher according to DtA statements.

<sup>3</sup> ERP stands for European Recovery Program.

ween assistance and ‘cash and carry’ effects. Secondly, these studies compare the supported firms with those that received no assistance at all without controlling for differences in both groups relating to important firm characteristics (firm size, legal form, economic sector etc.) that may influence the probability of receiving assistance and the propensity to grow.

In the context of this study the mean program impact is the potential difference in the average annual growth rates between the firms that firstly and exclusively received DtA assistance within two years after start-up and those that received no assistance at all. The analysis is medium-term since the growth rate of assisted and control group firms is examined over a six year interval. Therefore, ‘cash and carry’ effects should not play an important role. The empirical part of the paper then uses an econometric selection approach to estimate the mean program impact resulting from the receipt of public DtA start-up assistance. This approach controls for two potential selection sources. The first one concerns the fact whether or not the firms received public start-up assistance. The second one controls for the availability of suitable information to carry out the success analysis. This is necessary since only a part of the initially 12,000 available firm observations remains for the success analysis.

## **2 Data**

### **2.1 The DtA data base**

The DtA as the most important financier in the ‘Entrepreneurship/ Start-up’ segment provides more than 80 per cent of the German funds allocated in the 1990s. Major instruments are the ERP equity capital assistance program as well as the ERP and the DtA business start-up program. The ‘Entrepreneurship/Start-up’ segment contributed about 95 per cent of all

acceptances of assistance (about 460,000) and more than 65 per cent of the funds allocated (about 26 billions EURO) with its three major programs of all DtA assistance cases from 1990 until 1995 (DtA 2000). The funds are not provided directly by the DtA, but via the (local) house bank. Terms and conditions of the programs under evaluation change over time and are therefore not stated here.<sup>4</sup>

The DtA provides information in its data base to all acceptances of public assistance it made from 1990 until 1999. This data base contains about 776,000 data entries.<sup>5</sup> Rejected applications are not included, i.e. no information regarding the number of firms that applied for assistance but did not receive funding is available. Hence, no statements regarding the reasons for a rejection can be made. However, rejection of applications by the DtA is of minor relevance since the number of rejections ranges between 1 and 9 per cent (see appendix in Almus and Prantl 2001). Each data entry comprises information regarding the applying person (name, age, sex, address etc.) or applying firm (name, address) and the entity which is to be funded. This information will be used to identify the firms from the ZEW Entrepreneurship Study. Moreover, many details regarding terms and conditions of the investment cases is available. This information is used to select the cases in the programs of interest in the 'Entrepreneurship/Start-up' funding segment. Especially, the year of funding and the program indicator were necessary to identify the cases that received assistance at the right time in the right scheme. Moreover, information regarding the legal form and industry classification that were available in a revised version of the data base helped to identify the supported firms.

---

<sup>4</sup> Further information can be obtained from the DtA homepage (<http://www.dta.de>).

<sup>5</sup> For a more comprehensive description see Almus and Prantl (2001).

## 2.2 The ZEW Entrepreneurship Study

The second data set used is the ZEW Entrepreneurship Study and this one contains 12,000 observations that have been included in a telephone survey in 1999 (Almus et al. 2001).<sup>6</sup> Firms from the ZEW Foundation Panels East and West form the parent population for this survey (Almus et al. 2000). All firms considered when drawing the sample were founded between 1990 and 1993 according to CREDITREFORM<sup>7</sup> information, operate in the manufacturing, construction, trade or selected service sector branches and do not have the legal forms of freelance, registered society or registered cooperative. First of all, a stratification is made between Eastern and Western German firms. 6,000 firms each build the sample. Secondly, the other main stratification criterion is an indicator that gives information as to whether the firm has possibly exited the market or is still active.<sup>8</sup> Likewise, we oversample firms with these indicators when drawing the sample to get a satisfactory number of interviews with exiting firms for the empirical analysis, since these firm groups show a worse response pattern compared to firms which are still market active (Almus et al. 2001).

The main goal of the telephone survey was to gather information relating to annual employment development as well as the survival status for each of the 12,000 firms. We used a CATI (**C**omputer **A**ssisted **T**elephone **I**nterviewing) system to carry out the survey that was divided into three stages. In the first stage, a firm representative was expected to answer the survey questions. Afterwards, a current or former<sup>9</sup> firm participant (owner, manager, partner with the highest share) was chosen for all firms that did not answer the survey completely in the first stage. If no contact could be established with the selected firm participant or the person refused to answer the survey questions,

---

<sup>6</sup> The questionnaire is part of a project co-financed by the German National Science Foundation (DFG) under the grant LE1147/1-1.

<sup>7</sup> CREDITREFORM is the largest German credit rating agency.

<sup>8</sup> This indicator is based on CREDITREFORM information and points to problems in handling the data set or to different stages of bankruptcy proceedings.

<sup>9</sup> This mainly concerns the potential market exits.

a second owner/manager was chosen. With 3,702 complete interviews<sup>10</sup> out of 12,000 sample observations, the response rate was about 31 per cent, which is relatively high compared to other German studies conducted by telephone.<sup>11</sup>

## 2.3 Merging ZEW and DtA data

### 2.3.1 Preparing the data sets

The two data bases have to be merged to construct the data set that contains information on public start-up assistance activities. In the DtA data base the applicant (individual or firm) for start-up assistance defines a data entry. Firms and additionally all owners and/or participants represent the entities in the ZEW data. Up to this point the data bases contain many superfluous entries that have to be corrected.

Several entries in the DtA data base refer to the same object of investment, since the same person or firm can apply for the same assistance scheme in subsequent years or for various assistance schemes in the same year. Moreover, the same individual can apply for assistance in several programs and different years for various objects of investments. Therefore, observations that point to the same object of investment, but do not differ in their information content with respect to applicant's name, date of birth (only if a person applies for assistance) and address (street, zip code, place of residence) appear after the correction only once. This finally leads to about 500,000 data entries that differ in their information with respect to name, date of birth and address.<sup>12</sup>

---

<sup>10</sup> The term "complete interview" refers to a final status of the CATI system. Nevertheless the individual observations differ in their information content.

<sup>11</sup> Several firms refused to answer all questions but at least gave information as to whether they had exited the market or not. For analyses dealing with the survival of firms, a statement regarding the survival status is possible for an additional 2,234 firms.

<sup>12</sup> However, it is possible to trace the information back to all data sets, since this analysis focuses on assistance within two years after start-up. These 500,000 entries were further separated in two groups. The first one contains all acceptances of assistance from 1990 until 1995 since the primary focus is on tracking down recipients of start-up assistance within two years after foundation. The second group contains all acceptances of assistance from 1996 until 1999. All firms in the empirical analysis should not have received assistance in these years since the mean program effect of start-up assistance is the interesting variable.

The up to this point available information for the 12,000 observations from the ZEW Entrepreneurship Study cannot be used to link the two data bases. Problems arise because normally people and not firms apply for assistance. However, the ZEW Entrepreneurship Study is firm based. For this reason, the creation of additional data entries is necessary. Firstly, additional entries appear whenever changes in the firm name or firm address occur. These entries are necessary to find firms in the DtA data base that applied for assistance and the investment object is identical with the firm. Secondly, the ZEW Foundation Panels East and West as the parent populations of the ZEW Entrepreneurship Study contain details on both firms and owners/managers. Therefore, all available entries for any of the owners or managers are used in the merging process. A new entry appears if name or address details of the respective person change. These entries are necessary to find people in the DtA data set that applied for public assistance and the address of the applying person coincides with that of the investment object. Finally, we link all available name and date of birth information of firm participants with the respective firm addresses to find those entries in the DtA data base where firm participants applied for public assistance but stated the firm address in the application form. This huge amount of information is corrected for identical data entries. In the end about 200,000 data entries remain.

### **2.3.2 Merging**

The applied computer assisted heuristic string search approach is based on the comparison of values of the five variables (search indicators) name, date of birth, street, zip code and place of residence in the two data sets (see Almus and Prantl 2001) and the appendix in Janz et al. (2001) for detailed information). The indicators enter the search procedure with different weights that add up to 100 per cent. With respect to the degree of conformity the individual variables can reach the following maximum scores (name

of the firm or participant [40 per cent], date of birth [25 per cent]<sup>13</sup>, place of residence [13 per cent], zip code [12 per cent] and street [10 per cent]).

The approach then compares every entry in the DtA data with each ZEW entry and accordingly assigns a degree of conformity ranging from 0 to 100 per cent. After this time-consuming procedure, the results had to be separated into relevant and irrelevant merges. Matches with a degree of conformity of 78 per cent and more represent cases where no further manual control took place since at least the name, date of birth and two of the three address variables coincide. Cases with a degree of conformity below 50 per cent are irrelevant and will not belong to the treatment group. The remaining cases with values between 50 and 77 per cent had to be checked manually.

Finally 2,552 out of the 12,000 initial observations from the ZEW Entrepreneurship Study received DtA assistance between 1990 and 1995 and are therefore eligible to belong to the potential treatment group. The remaining 9,448 did not receive DtA assistance between 1990 and 1995 and build up the temporary potential control group. Up to this point it is possible that firms from the treatment and control group might have received assistance between 1996 and 1999.

Furthermore, several corrections took place that scaled down the number of firm in the potential treatment and control group.

- Excluded are firms that received start-up assistance between 1996 and 1999 from both the treatment and the potential control group.
- There are entries in the DtA data base that point to different firms in the ZEW Entrepreneurship Study. These firms were excluded both from the group of supported firms and the potential control group.

---

<sup>13</sup> The date of birth is not available if a firm applies for public assistance. In these cases the maximum score of conformity only can reach 75 per cent. Further search variables (legal form, economic sector classification) were not available in the DtA data set at the time the merging took place.

- Firms that received assistance between 1990 and 1995 but not in the first two years after start-up neither entered the group of assisted start-ups nor the group of comparisons.
- Firms that received public assistance in DtA schemes other than the ERP equity capital assistance program and the ERP or DtA business start-up program neither entered the group of assisted start-ups nor the group of comparisons. This check was necessary since the DtA data base contains assistance acceptances for all programs administered by the DtA.

Taking these points together, only 1,726 ( $N_S$ ) observations with valid interview information remain (see section 5 for more details). 472 ( $N_1$ ) of them have received assistance at least once within the first two years after start-up. The remaining 1,254 firms ( $N_0$ ) that receive no public assistance at all represent the potential control group for the subset of observations with satisfactory interview information. This separation is also possible for firms without necessary interview information. 1,472 firms exclusively received DtA assistance within the first two years after start-up and the remaining 7,629 firms did not.<sup>14</sup>

## 3 Evaluation

### 3.1 The basic problem

An evaluation can take place either in experimental or non-experimental situations. In experimental situations all individuals eligible for the measure or program have a priori the same probability to belong to the group of participants or to the control group. Hence, the assignment to the program or measure is random. Non-experimental program designs like the one in this study

---

<sup>14</sup> Note that the figures did not add up to 12,000. This is due to the loss of observation with missing information in important exogenous variables (see section 4 for more details). This separation is necessary since I control for two potential selection sources in empirical analysis.

cannot assure a random assignment since it is likely that individual with particular observable and unobservable characteristics (experience, motivation, information advantages) have a higher probability to enter the program or measure. And this non-coincidental nature of the experiment has to be controlled for. This might be interpreted as a disadvantage of non-experimental settings. However, there is no a priori superiority of experimental against non-experimental data according to Heckman and Smith (1995).

Hence, we assume a non-experimental setting where the sample is separated into participants in a program or measure (treatment group) and non-participants (potential control group). Furthermore, we observe a success variable  $Y$  ( $Y^1$  for treatment group observations and  $Y^0$  for control group observations). The evaluation then aims to calculate the mean impact  $\theta^1$  of the program for the participants (Heckman and Hotz 1989):

$$\theta^1 = E_X[\theta^1(X)|D = 1] := E_X[E(Y^1 - Y^0|X, D = 1)|D = 1] \quad (1)$$

Here,  $D$  is an indicator variable that either signals if individual  $i$  has taken part in the measure or program ( $D_i = 1$ ) or not ( $D_i = 0$ ),  $E_X[\bullet]$  is the iterated expected value and  $X$  is a vector of exogenous variables.

It turns out that there is an identification problem since the term  $E(Y^0|X, D = 1)$  in equation (1) is by definition not observable. This leads to a potential selection bias  $\eta(X)$

$$\begin{aligned} \theta^1 &= E(Y^1|X, D = 1) - E(Y^0|X, D = 0) \\ &+ E(Y^0|X, D = 0) - E(Y^0|X, D = 1) \\ \theta^1 &= E(Y^1|X, D = 1) - E(Y^0|X, D = 0) + \eta(X) \end{aligned} \quad (2)$$

There are several sources of selection when evaluating the success of a program and you must be aware of selection biases whenever the available data do not represent a random sample from the parent population, i.e. individuals, persons, firms etc. do not have equal chances to enter the random sample

that is examined (Entorf 2000). Generally, a selection bias may result from self selection (decision by the individual itself), administrative selection rules and/or decisions by the individuals who collect the data. These can be summarized in the case of the evaluation of public start-up assistance on behalf of the DtA as

- entrepreneurial self selection (i.e. amount of information the entrepreneur possesses and access to information that the entrepreneur needs, the extent to which the entrepreneur needs external capital and the attitude toward outside inference (giving away part of the control)),
- selection on behalf of the (local) house bank (i.e. decisions on the bank-customer-relationship; decisions on the financing structure of the firm) and
- institutional selection on the part of the DtA (i.e. program requirements that rule out the participation of firms with special characteristics; incentive of the DtA to maximize the program success)

The first two terms in equation (2) can be estimated unbiased by calculating the arithmetic mean of the two sub-populations. However, neither  $E(Y^0|X, D = 1)$  nor  $\eta(X)$  are identified. No selection bias, i.e.  $\eta(X)$  equals zero, would occur if there were no differences between assisted and non-assisted start-ups in terms of their success or performance prior to the receipt of assistance (Pfeiffer and Reize 2000). If this cannot be assured, the mean program impact will be biased due to the classical selection bias (Heckman 1974). Moreover, that part of the selection bias which is due to unobservable variables influencing both the success and the receipt of public start-up assistance might itself remain unobserved (Heckman and Hotz 1989). This led to the development of approaches that can be used to estimate the mean program impact in non-experimental studies and that try to control for selection biases. The three ones most often applied will be briefly presented in the following.

- The “difference-in-differences” method (Ashenfelter 1978, Ashenfelter and Card 1985) became popular with the availability of panel data sets. Here, potential selection biases stemming from observable time invariant variables vanish in the linear model if differences are calculated over time (Fitzenberger and Prey 1998).
- Matching methods became more and more popular in the evaluation of labor market programs and are based on the model of potential outcomes. Rosenbaum and Rubin (1983) point out that matching “[. . .] is a method for selecting units from a large reservoir of potential comparisons to produce a comparison group of modest size in which the distribution of covariates is similar to the distribution in the treated group.” Therefore, we seek a ‘perfect twin’ for each observation of the treatment group, i.e. at least one observation of the potential control group that is as similar as possible to the treated observation with respect to a given distance measure. The success of these approaches depends on several conditions that allow the identification of the potential effect (Heckman et al. 1998a, Heckman et al. 1999).
- Complete econometric selection models simultaneously estimate participation in and success of the program or measure. These models depend on restrictive assumptions regarding the error terms and their distribution that often cannot be interpreted economically. Therefore, these models have often been criticized (Ashenfelter and Card 1985). However, Heckman and Hotz (1989) point out that the application of parametric models leads to satisfying results. Parametric instrument variable estimators have increasingly gained attention in the last few years may be seen as variant of this class of estimators (Angrist et al. 1996).

All these approaches have their advantages and disadvantages and there are actually no guidelines as to when to use statistical matching or econometric evaluation models. “[...] Thus the choice of an appropriate econometric model critically depends on the data on which it is applied” (Heckman et al. 1998b).

Moreover, Heckman and Hotz (1989) conclude that “[...] there is no objective way to choose among alternative nonexperimental estimators.” I finally apply a parametric approach since two potential selection sources, i.e. the receipt of assistance as well as the availability of sufficient information in the ZEW Entrepreneurship Study for the respective firms should be controlled for.

## 4 Descriptives

Though the telephone survey did not include all 12,000 firms, I have comprehensive information from the ZEW Foundation Panels East and West for all observations of the ZEW Entrepreneurship Study (see Almus et al. (2001) for more details). Unfortunately, not all of the 12,000 firms can be used for the empirical analysis. The most important reasons for the exclusion of observations are

- for several firms in the ZEW Entrepreneurship Study the number of employees at start-up recorded by CREDITREFORM is not available due to the design of the random sample (Almus et al. 2001),
- firms received support in programs other than those under evaluation and may not enter the treatment or potential control group for that reason,
- firms received start-up assistance more than two years after start-up and may therefore not enter the treatment or potential control group.

Finally, 10,827 observations remain. Table 1 contains descriptive statistics for the relevant variables used in the empirical analysis. This comprehensive data set is used to control for potential selection sources in the data mentioned before. About 16 per cent of the firms have suitable interview information for the analysis to calculate the mean program impact. 18 per cent of the firms have received public start-up assistance within the first two years of existence but not later on. The remaining firms did not receive any DtA assistance

at all. Finally, the average number of employees at start-up according to CREDITREFORM information is 19 whereas the median is two.<sup>15</sup>

>> *insert Table 1 about here* <<

The growth estimation to assess the mean program impact only considers observations that fulfill several conditions. The firms used were established between 1990 and 1993 according to the survey question in the ZEW Entrepreneurship Study. This leads to the exclusion of some firms from the data set which either have foundation dates before 1990 or after 1993 as reported in the telephone interview. Moreover, I exclude firms that represent partial or complete take-overs because this study focuses on the mean program impact of public start-up assistance on the subsequent employment growth of new, original start-ups. Both surviving firms and market exits must have a valid number of employees ( $E$ ) in the start-up year to be included in the empirical analysis. The firms also possess an employment number in the sixth year after start-up.<sup>16</sup> Only market exits that do not survive the fifth year after start-up do not have a second employment number. The number of employees is set to zero in the year after the market exit for these firms. The discrete average annual growth rate  $Y_i$  for the  $i$ -th firm is then calculated as

$$Y_i = \frac{E_{i,\tau} - E_{i,t}}{(\tau - t) E_{i,t}} . \quad (3)$$

The index  $t$  characterizes the start-up year and  $\tau$  represents the sixth year after start-up for all firms with two valid employment numbers or the year after the market exit for the remaining firms.<sup>17</sup> Finally, firms are not allowed

<sup>15</sup> This large average number of employees mainly results from partial or complete take-overs that are included here but not in the final estimation to assess the success of the measure. Reasons why observations with such high employment figures enter the ZEW Foundation Panels East and West as well as the ZEW Entrepreneurship Study are explained in Almus et al. (2001) or Almus et al. (2000).

<sup>16</sup> A 1990 start-up must therefore have a valid number of employees for 1990 and 1996. I use the six year interval since the latest start-up year is 1993 and the latest year where the number of employees was asked is 1999. Therefore, the maximum time span available for all firms is six years.

<sup>17</sup> The calculation of the continuous growth rate  $(\ln E_{i,\tau} - \ln E_{i,t})/(\tau - t)$  is not possible for the market exits, since here is  $E_{i,\tau} = 0$ . This would lead to a loss of observations. Therefore, the discrete growth measure is used.

to have missing values for the exogenous variables used in the growth analysis. 1,726 observations remain after these correction and Table 2 contains the descriptive statistics.

>> *insert Table 2 about here* <<

About 27 per cent of the firms have exclusively received assistance within the first two years after start-up. According to Table 1 this number increases by about 9 percentage points. Furthermore, the share of Western German firms decreases to 42 per cent. The number of employees at start-up now only amounts to about 7 which is a more realistic figure for start-ups. Finally, two tailed t-tests were carried out to see whether the assisted and non-assisted firms differ significantly in observable characteristics. This concerns nearly all variables in the data set and indicates that public assistance on part of the DtA does not follow the principle of ‘equal shares for all’. The group of firms that received public assistance is rather an exclusive one that shows special characteristics.

## 5 Econometrics and results

The empirical approach has to take into account the selection bias in equation (2) when estimating the mean program impact. The comprehensive data set at hand can help to solve this problem, or at least to minimize the bias that is due to observable characteristics. However, I cannot totally rule out the occurrence of a selection bias that is due to unobservables. Despite the huge information content in the data set, there are unobservable determinants that influence the success and the receipt of assistance. Moreover, a second source of selection exists that might arise due to the fact that the firms with complete interview information are not a random sample of all firms from the ZEW Entrepreneurship Study. And only these 1,726 observations with complete interview information enter the econometric analysis to assess the mean program effect. Hence, it might be possible that firms taking part in

the survey have better growth chances regardless of whether they took part in the survey (self selection).

The trivariate structure of the following econometric model will help to minimize these possible selection biases and to assess the mean program impact unbiased (Pfeiffer and Reize 2000). Suppose, there is a latent variable  $S_i^*$  that can be observed for all  $N(= 10, 827)$ -firms with valid and suitable information (see section 4 for details). This variable indicates whether firm  $i$  has complete interview information ( $S_i = 1$ ) or not ( $S_i = 0$ )

$$\begin{aligned} S_i^* &= X'_{Si} \beta_S + D_i \alpha + \varepsilon_{Si} \quad \forall i = 1, \dots, N & (4) \\ S_i &= 1 \text{ iff } S_i^* > 0 \\ S_i &= 0 \text{ iff } S_i^* \leq 0 . \end{aligned}$$

The vector  $X_{Si}$  contains exogenous variables that may influence the probability of having complete interviews and  $D_i$  indicates whether firm  $i$  received assistance or not. I include this variable since assistance may have an impact on the survival of firms and therefore on the probability of taking part in the survey that resulted in the ZEW Entrepreneurship Study. The interesting parameters are  $\beta_S$  and  $\alpha$ . Information also exists for each of the  $N$  sample firms if they have exclusively received DtA assistance in the first two years after start-up ( $D_i = 1$ ) or not ( $D_i = 0$ ). This results in the second equation

$$\begin{aligned} D_i^* &= X'_{Di} \beta_D + \varepsilon_{Di} \quad \forall i = 1, \dots, N & (5) \\ D_i &= 1 \text{ iff } D_i^* > 0 \\ D_i &= 0 \text{ iff } D_i^* \leq 0 \end{aligned}$$

where  $X_{Di}$  and  $\beta_D$  have analog functions as  $X_{Si}$  and  $\beta_S$  in equation (4). This system of equations (4) and (5) allows for endogeneity of the receipt of assistance. The third equation models the success of the firms, i.e. the average annual employment growth rate

$$\begin{aligned} Y_i &= X'_{Yi} \beta_Y + D_i \theta^1 + \varepsilon_{Yi} \quad \forall i = 1, \dots, N_S & (6) \\ Y_i &= D_i Y_i^1 + (1 - D_i) Y_i^0 , \end{aligned}$$

but is only defined for the  $N_S$  observations with complete interviews ( $S_i = 1$ ).  $Y_i^1(Y_i^0)$  indicates the outcome (average annual growth rate) in cases where the firm received (did not receive) public start-up assistance. The vector  $X_{Y_i}$  contains exogenous variables that may influence the growth rate of firms and  $\beta_Y$  is the interesting parameter vector. As in equation (4),  $D_i$  indicates whether firm  $i$  has received start-up assistance or not and  $\theta^1$  is the interesting parameter that measures the mean program impact.

The error terms  $\varepsilon_{S_i}$  and  $\varepsilon_{D_i}$  in equations (4) and (5) are assumed to be normally distributed with mean 0 and variance 1. The error term in the growth equation is assumed to be normally distributed with mean 0 and variance  $\sigma_Y^2$ . Heckman and Hotz (1989) and Angrist et al. (1996) point out that a dependence between  $D_i$  and  $\varepsilon_{Y_i}$  in equation (6) can arise for two reasons. Firstly, there may be a dependence between  $X_{D_i}$  and  $\varepsilon_{Y_i}$  that is called selection on observables. Secondly, the correlation between  $\varepsilon_{D_i}$  and  $\varepsilon_{Y_i}$  is called selection on unobservables and concerns the questions as to whether firms that received public start-up assistance would have better growth prospects whether or not they received such funding (self selection). Additionally, the model structure can generate a further source of selection on unobservables, i.e. a potential correlation between  $\varepsilon_{S_i}$  and  $\varepsilon_{Y_i}$  that is measured with  $\rho$ .

This trivariate model structure then has the following variance-covariance-matrix  $\Sigma$

$$\Sigma = Var \begin{pmatrix} \varepsilon_{Y_i} \\ \varepsilon_{D_i} \\ \varepsilon_{S_i} \end{pmatrix} = \begin{pmatrix} \sigma_Y^2 & \sigma_{YD} & \sigma_{YS} \\ \sigma_{YD} & 1 & \sigma_{DS} \\ \sigma_{YS} & \sigma_{DS} & 1 \end{pmatrix} .$$

Note that the model is identified as long as the vector  $X_{D_i}$  includes at least one variable that is neither included in  $X_{S_i}$  nor in  $X_{Y_i}$ , and the error terms fulfill the distributional assumptions (Heckman et al. 1999, Puhani 2000). Hence, the model can be estimated with the following two-step approach

(Pfeiffer and Reize 2000). The first step consists of estimating a bivariate probit to simultaneously determine the probability of having complete interview information (equation (4)) and receiving public start-up assistance (equation (5)). Besides the bivariate structure I have to take the endogeneity of  $D_i$  in equation (4) into account. Hence, I estimate the two equations simultaneously (Mallar 1977, Gourieroux 2000). Initially, I estimate equation (5) and the reduced form of equation (4)

$$S_i^* = X'_{Si} \beta_S + \alpha(X'_{Di} \beta_D + \varepsilon_{Di}) + \varepsilon_{Si}$$

using a bivariate probit model. This results in an estimate for  $E[D_i|X_{Si}, X_{Di}]$ . Afterwards this estimated value is inserted in equation (4) and the system of equations (4) and (5) is once again estimated using a bivariate probit model. Though this estimation leads to consistent parameter estimates, there are problems with the estimated standard errors that have to be corrected (Murphy and Topel 1985, Greene 1998). This is due to the fact that the estimated regressor  $E[D_i|X_{Si}, X_{Di}]$  is measured with sampling error. To overcome this problem, I calculate bootstrapped standard errors to carry out inferences (Efron and Tibshirani 1993). Table 3 contains the results of the bivariate probit model.

>> *insert Table 3 about here* <<

The interpretation of the results of the bivariate probit model concentrates on the one hand on the potential effect of receipt of start-up assistance on the probability of having complete interview information in the survey. On the other hand the correlation  $\rho$  between the error terms  $\varepsilon_{Si}$  and  $\varepsilon_{Di}$  is a parameter of interest. It turns out that the receipt of assistance significantly increases the probability of taking part in the survey leading to usable interview information for the success analysis. This may be due to the fact that firms that receive DtA start-up assistance have better survival chances (Almus and Prantl 2001) and for that reason have the chance to take part in the interview. Moreover, firms that apply for DtA assistance may have

better information and therefore are more likely to take part in the interview. Additionally, there is a significant correlation between the error terms  $\varepsilon_{S_i}$  and  $\varepsilon_{D_i}$  that indicates potential selection on unobservables that has to be controlled for. The correlation coefficient  $\rho$  amounts to a value of about 0.15 (significant at any conventional level).

The next step in the estimation procedure consists of calculating two inverse Mill's ratios  $\lambda_S$  and  $\lambda_D$  are calculated that enter the second stage estimation as additional regressors. One should keep in mind that the growth estimation includes all  $N_S$ -firms with complete interviews, whether or not they received public start-up assistance. Therefore, the Mill's ratios have the following form (Tunali 1986, Goux and Maurin 2000)

$$\lambda_{D_i} = \begin{cases} \phi(Z_{D_i}) \frac{\Phi((Z_{S_i} - \rho Z_{D_i}) / (1 - \rho^2)^{1/2})}{\Phi_2(Z_{D_i}, Z_{S_i}, \rho)} & \text{if } D_i = 1 \\ -\phi(Z_{D_i}) \frac{\Phi((Z_{S_i} - \rho Z_{D_i}) / (1 - \rho^2)^{1/2})}{\Phi_2(-Z_{D_i}, Z_{S_i}, -\rho)} & \text{if } D_i = 0 \end{cases} \quad (7)$$

$$\lambda_{S_i} = \begin{cases} \phi(Z_{S_i}) \frac{\Phi((Z_{D_i} - \rho Z_{S_i}) / (1 - \rho^2)^{1/2})}{\Phi_2(Z_{D_i}, Z_{S_i}, \rho)} & \text{if } D_i = 1 \\ \phi(Z_{S_i}) \frac{\Phi((-Z_{D_i} + \rho Z_{S_i}) / (1 - \rho^2)^{1/2})}{\Phi_2(-Z_{D_i}, Z_{S_i}, -\rho)} & \text{if } D_i = 0 \end{cases} \quad (8)$$

In equations (7) and (8) stand

$$\begin{aligned} Z_{S_i} &= X'_{S_i} \beta_S + D_i \alpha \\ Z_{D_i} &= X'_{D_i} \beta_D \quad , \end{aligned}$$

$\phi(\bullet)$  resp  $\Phi(\bullet)$  represent the univariate cdf and pdf of the standard normal distribution,  $\Phi_2(\bullet)$  stays for the bivariate pdf of the standard normal distribution and  $\rho$  is the correlation coefficient between  $\varepsilon_{D_i}$  and  $\varepsilon_{S_i}$ . An OLS estimation based on equation (6) augmented with the two estimated inverse Mill's ratios ( $\hat{\lambda}_{D_i}$  and  $\hat{\lambda}_{S_i}$ ) is used to evaluate the success, i.e. the mean program impact  $\theta^1$ , of start-up assistance programs administered by the DtA

$$Y_i = X'_{Y_i} \beta_Y + D_i \theta^1 + \gamma_S \hat{\lambda}_{S_i} + \gamma_D \hat{\lambda}_{D_i} + u_{Y_i} \quad \forall i = 1, \dots, N_1 \quad (9)$$

The inclusion of the two estimated Mill's ratios ensures that the error term  $u_{Y_i}$  is normally distributed with mean 0 and variance  $\sigma_u^2$  as long as the

distributional assumptions are fulfilled. The model is identified as long as the vector  $X_{Di}$  includes at least one variable with significant impact that is neither included in  $X_{Si}$  nor in  $X_{Yi}$  as mentioned above. The variable “bank customer density in 1990/91” that measures the number of inhabitants (in 1,000) per bank sub-office in 1990 for Western German counties and in 1991 for Eastern German counties serves as identifying variable since it has a significant impact (see Table 3). Then, the OLS estimation provides unbiased results for  $\beta_Y$  and  $\theta^1$  as well as for the parameters  $\gamma_S$  and  $\gamma_D$ . These two parameters control for potential correlations between the error terms in equations (4) and (5) on the one hand and (6) on the other. The inclusion of the two estimated Mill’s ratios ( $\hat{\lambda}_{Di}$  and  $\hat{\lambda}_{Si}$ ) as additional exogenous variables leads once again to inconsistent estimated standard errors. The same reason as in the bivariate probit model applies. To obtain consistent standard errors the derivatives of the Mill’s ratios must be computed which is very cumbersome (Reize 2001). Therefore, I once again estimate and use bootstrapped standard errors.

The results, which are not reported here but are available on request, indicate that only  $\hat{\lambda}_{Si}$  has a significant impact whereas  $\hat{\lambda}_{Di}$  is statistically insignificant. This simplifies the model in equation (9) since controlling for unobserved heterogeneity between the selection process of assisted firms and employment growth (correlation between  $\varepsilon_{Yi}$  and  $\varepsilon_{Di}$ ) is not necessary. Therefore only  $\hat{\lambda}_{Si}$  serves as additional exogenous variable in the growth equation to assess the mean program impact of public start-up assistance. Table 4 contains the results of the OLS estimation. As mentioned before bootstrapped standard error are used to make inferences.

>> *insert Table 4 about here* <<

The goodness of fit measure of the estimation of the mean program impact is the adjusted  $R^2$ . This measure is about 6 per cent, i.e. about 6 per cent of the variance of the growth rate ( $Y_i$ ) can be explained with the model at

hand. The mean program impact amounts to a value of 0.07. The effect is statistically different from zero at the 5 per cent level of significance. This means that other things equal (i.e. same set of observable characteristics) firms realize on average a 7.0 percentage point higher growth rate if they have exclusively received public start-up assistance in the first two years after start-up. Hence, start-up assistance programs administered by the DtA have a significantly positive medium-term impact on the firm development measured with employment growth. This confirms the results of Almus and Prantl (2001) who found that start-ups receiving DtA assistance over their entire life time have a higher probability of survival and higher employment growth rates in case of survival. However, the amount of the effect between both studies differs. This is mainly due to the application of different approaches, since Almus and Prantl (2001) use a non-parametric matching approach that tries to find matched pairs ('perfect twins') of assisted and non-assisted firms that do not differ in important characteristics measured at start-up. Moreover, receipt of assistance is in Almus and Prantl (2001) not restricted to the first two years after start-up. Finally, partial and complete take-overs that have a start-up date between 1990 and 1993 were included besides original start-ups. The results contradict the findings of other studies that use a similar approach like in this analysis (Pfeiffer and Reize 2000, Reize 2001). These two studies find an insignificant influence of start-up promotion for unemployed entrepreneurs on the growth of their firms or businesses. However, the potential firm founders are not comparable to the ones in my study (unemployed entrepreneurs vs all entrepreneurs that received DtA assistance).

Self selection that is controlled for by the variable  $\hat{\lambda}_{S_i}$  has a significant impact, i.e. there is a selection on unobservables resulting from a correlation between the error terms  $\varepsilon_{Y_i}$  and  $\varepsilon_{S_i}$ . Hence, unobservable characteristics that determine the availability of sufficient interview information have a significant impact on employment growth. As mentioned above there is no selection

bias observable coming from a correlation between the error terms  $\varepsilon_{Y_i}$  and  $\varepsilon_{D_i}$ . There is seemingly no growth differential that results from unobservable factors that determine whether or not a firm received public assistance.

The effects of the remaining variables are mostly as expected and will be summarized only briefly. Initial firm size has a negative influence on the average annual growth rate indicating a deviation from Gibrat's law and hence a higher growth potential for small firms (Sutton 1997). Firms with limited-liability legal forms (GmbH [non-public limited liability firms], GmbH&Co.KG [commercial partnerships formed with a non-public limited liability firms] or Aktiengesellschaft [joint-stock company]) achieve *ceteris paribus* higher growth rates than firms with remaining legal forms. Firms with limited-liability legal forms have higher incentives to pursue more risky projects that lead in the case of success among others to higher employment growth rates (Stiglitz and Weiss 1981). Firms founded in 1992 have worse employment growth chances compared to firms from the remaining start-up cohorts. Western German firms achieve on average lower growth rates than their Eastern German counterparts. These effects as well as the effects coming from the economic sectors are due to the special situation in Eastern Germany in the first years after reunification (Almus 2000). Firms from manufacturing branches, construction, and business related services achieve other things equal higher growth rates compared to firms from the reference sector trade.

Moreover, the average growth rate is not altered if other firms hold a share on the firms examined. The availability of knowledge and financial support as well as connections to suppliers and/or customers does not seem to favor the development of these firms. Entrepreneurial teams have no advantage with respect to the employment performance compared to start-ups where only one person was involved. Finally, the human capital of the owners/managers has an influence on the growth rate. Firms with owners and/or managers that are professors or have a doctorate achieve on average higher employ-

ment growth rates.

## 6 Summary

Every year The DtA spends billions of Euro on assisting new firms and entrepreneurs in its ‘Entrepreneurship/Start-up’ segment. The focus of this study is therefore to examine whether these programs have a medium-term effect on the employment development of new firms. I use an econometric selection approach to estimate the mean impact of start-up assistance programs administered by the DtA. With the ZEW Entrepreneurship Study I have a comprehensive data set to conduct the empirical analysis. Using DtA data that contain all assistance acceptances from 1990 until 1999, I can ascertain whether a firm from the ZEW Entrepreneurship Study has exclusively received start-up assistance administered by the DtA. The estimation results show that the receipt of assistance significantly increases the average employment growth rates of firms over a six year period. Start-up assistance programs of the DtA significantly improve the employment performance of the assisted firms under consideration on average by seven percentage points.

An important point to be mentioned concerns the employment effects of the examined firms. The significantly higher average employment growth rates of the assisted firms are a necessary but not a sufficient argument for the employment creating effects of DtA assistance. It is not possible to observe indirect effects, i.e. potential crowding out effects, that result from the firms under investigation. It is possible that the employment created by these firms leads to a destruction of jobs in other firms, i.e. competitors. But the market entry of new firms would even lead in this case to a rejuvenation effect in the economy, since the examined firms are young (not older than 9 years). Moreover, the firms under consideration may exhibit increasing innovation activities and generate positive spill-over effects.

There are some points that limit the potential to make inferences from the results obtained. Only assistance from the DtA is observed. A control for financial support from alternative programs fails due to the availability of such data. Therefore, I assume that funds from alternative programs are randomly allocated. But this problem is of minor relevance. According to DtA statements the bank covered more than 80 per cent of public start-up assistance in the 1990s. Moreover, the number of support events as well as the amount of financial support may have an influence on the mean program impact but is not regarded yet. Up to now, DtA assistance is simply coded as a  $[0,1]$  decision (either the firm has received assistance or not). The results obtained are an important prerequisite for decision makers when they have to determine whether a firm should be supported or not. However, the analysis is based on average results and should therefore not be used as a single decision criterion as to whether a firm should be supported or not. Finally, the calculation of the average employment growth rates only considers the employment at start-up and in the sixth year after start-up. Hence, transitory employment fluctuations are not taken into account.

## Literatur

- Almus, M. (2000), What Characterizes a Fast Growing Firm?, Discussion Paper, No. 00-64, Centre for European Economic Research (ZEW), Mannheim.
- Almus, M., D. Engel and S. Prantl (2000), The "Mannheim Foundation Panels" of the Centre for European Economic Research (ZEW), Documentation, No. 00-02, Centre for European Economic Research (ZEW), Mannheim.
- Almus, M., S. Prantl, J. Brüderl, K. Stahl und M. Woywode (2001), Die ZEW-Gründerstudie – Konzeption und Erhebung, Documentation, No. 01-01, Centre for European Economic Research (ZEW), Mannheim.
- Almus, M. and S. Prantl (2001), Bessere Unternehmensentwicklung durch Gründungsförderung, *Wissenschaftliche Reihe der DtA*, Vol. 15, Deutsche Ausgleichsbank (DtA), Bonn.
- Angrist, J., G. Imbens and D.B. Rubin (1996), Identification of Causal Effects Using Instrumental Variables, *Journal of the American Statistical Association*, Vol. 91, pp. 444-455.
- Ashenfelter, O. (1978), Estimating the Effect of Training Programs on Earnings, *Review of Economics and Statistics*, Vol. 60, pp. 4-57.
- Ashenfelter, O. and D. Card (1985), Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs, *Review of Economics and Statistics*, Vol. 67, pp. 648-660.
- Battistin, E., A. Gavosto und E. Rettore (2001), Why do subsidized firms survive longer? An evaluation of a program promoting youth entrepreneurship in Italy, in Lechner, M. und F. Pfeiffer (eds.), *Econometric Evaluation of Labour Market Policies*, Physica, Heidelberg, pp. 153-181.
- Brüderl, J., P. Preisendörfer und R. Ziegler (1993), Staatliche Gründungsfinanzierung und der Erfolg neugegründeter Betriebe, *Jahrbücher für Nationalökonomie und Statistik*, Vol. 212, pp. 13-32.
- David, P.A., B.H. Hall und A.A. Toole (2000), Is public R&D a complement or substitute for private R&D? A review of the econometric evidence, *Research Policy*, Vol. 29, pp. 497-529.

- Dehejia, R.H. and S. Wahba (1999), Causal effects in nonexperimental studies: reevaluating the evaluation of training programs, *Journal of the American Statistical Association*, Vol. 94, pp. 1053-1062.
- DtA [Deutsche Ausgleichsbank](2000), DtA-Förderreport, 10 Jahre DtA-Förderung 1990-1999, Bonn.
- Efron, B. and R.S. Tibshirani (1993), *An Introduction to the Bootstrap*, New York, Chapman & Hall.
- Entorf, H. (2000), James Heckman and Daniel Mc Fadden: Nobelpreis für die Wegbereiter der Mikroökonomie, Discussion Paper, No. 26, University of Würzburg.
- Fitzenberger, B. and H. Prey (1998), Beschäftigungs- und Verdienstwirkungen von Weiterbildungsmaßnahmen im ostdeutschen Transformationsprozeß: Eine Methodenkritik, in Pfeiffer, F. and W. Pohlmeier (eds.), *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, Baden Baden, NOMOS, pp. 39-96.
- Gourieroux, C. (2000), *Econometrics of Qualitative Dependent Variables*, Cambridge, Cambridge University Press.
- Goux, D. and E. Maurin (2000), Returns to Firm-Provided Training: Evidence from French Worker-Firm Matched Data, *Labour Economics*, Vol. 7, pp. 1-19.
- Greene, W.H. (1998), Gender Economics in Courses in Liberal Arts Colleges: Further Results, *Journal of Economic Education*, Vol. 29, pp. 291-300.
- Harhoff, D. and T. Koerting (1998), Lending Relationship in Germany - Empirical Evidence from Germany, *Journal of Banking and Finance*, Vol. 22, pp. 1317-1353.
- Heckman, J.J. (1974), Shadow prices, market wages, and labor supply, *Econometrica*, Vol. 42, pp. 679-694.
- Heckman, J.J. and V.J. Hotz (1989), Choosing Among Nonexperimental Methods for the Estimating the Impact of Social Programs: The Case of Manpower Training, *Journal of the American Statistical Association*, Vol. 84, pp. 862-880.
- Heckman, J.J. and J.A. Smith (1995), Assessing the case for social experiments, *Journal of Economic Perspectives*, Vol. 9, pp. 85-110.

- Heckman, J.J. and J.A. Smith (1997), Making the most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts, *Review of Economic Studies*, Vol. 64, pp. 487-535.
- Heckman, J.J., H. Ichimura and P. Todd (1997), Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, *Review of Economic Studies*, Vol. 64, pp. 605-654.
- Heckman, J.J., H. Ichimura and P. Todd (1998a), Matching as an Econometric Evaluation Estimator, *Review of Economic Studies*, Vol. 65, pp. 261-294.
- Heckman, J.J., H. Ichimura, J. Smith and P. Todd (1998b), Characterizing Selection Bias using Experimental Data, *Econometrica*, Vol. 66, pp. 1017-1098.
- Heckman, J.J., R.J. LaLonde and J.A. Smith (1999), The economics and econometrics of active labor market programs, in Ashenfelter, O. and D. Card (eds.), *Handbook of labor economics*, Vol. 3a, Amsterdam, pp. 1865-2097.
- Janz, N., G. Licht and T. Doherr (2001), Innovation Activities and European Patenting of German Firms, mimeo, Centre for European Economic Research (ZEW), Mannheim.
- Klette, T.J., J. Møen und Z. Griliches (2000), Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies, *Research Policy*, Vol. 29, pp. 471-495.
- LaLonde, R.J. (1986), Evaluating the econometric evaluations of training programs with experimental data, *The American Economic Review*, Vol. 76, pp. 604-619.
- Lechner, M. (1999), Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification, *Journal of Business and Economic Statistics*, Vol. 17, pp. 74-90.
- Lerner, J. (1999), The Government as Venture Capitalist: the Long-Run Impact of the SBIR Program, *Journal of Business*, Vol. 72, pp. 285-318.
- Mallar, C.D. (1977), The Estimation of Simultaneous Probability Models, *Econometrica*, Vol. 45, pp. 1717-1722.

- Murphey, K.M. and R.H. Topel (1985), Estimation and Inference in Two-Step Econometric Models, *Journal of Business & Economic Statistics*, Vol. 3, pp. 370-379.
- Pfeiffer, F. and F. Reize (2000), Business start-ups by the unemployed – econometric analysis based on firm data, *Labour Economics*, Vol. 7, pp. 629-663.
- Puhani, P. (2000), The Heckman Correction for Sample Selection and Its Critique – A Short Survey, *Journal of Economic Surveys*, Vol. 14, pp. 53-68.
- Reize, F. (2001), FIML Estimation of a Bivariate Probit Selection Rule – An Application on Firm Growth and Subsidisation, Discussion Paper, No. 01-13, Centre for European Economic Research (ZEW), Mannheim.
- Rosenbaum, P.R. and D.B. Rubin (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika*, Vol. 70, pp. 41-55.
- Stiglitz, J.E. and A. Weiss (1981) Credit rationing in markets with imperfect information, *The American Economic Review*, Vol. 71, pp. 393-410.
- Sutton, J. (1997) Gibrats' s legacy, *Journal of Economic Literature*, Vol. 35, pp. 40-59.
- Tunali, I. (1986) A General Structure for Models of Double-Selection and an Application to a Joint Migration/Earning Process with Reimmigration, *Research in Labor Economics*, Vol. 8, pp. 235-282.
- Wallsten, S.J. (2000), The Effects of Government–Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research Program, *RAND Journal of Economics*, Vol. 31(1), pp. 82-100.

# Appendix

**Tabelle 1:** Descriptive Statistics for the observations used in the bivariate probit model

variable	mean/share	STDV
interview available	0.159	0.366
assisted firms	0.180	0.384
number of employees at start-up	19.522	317.839
limited liability legal form	0.449	0.497
start-up in 1990	0.286	0.452
start-up in 1991	0.259	0.438
start-up in 1992	0.214	0.410
start-up in 1993	0.241	0.428
one firm owner	0.260	0.439
other firm(s) involved in start-up	0.105	0.307
very high human capital <sup>a)</sup>	0.035	0.185
high human capital <sup>a)</sup>	0.329	0.470
medium human capital <sup>a)</sup>	0.357	0.479
low human capital <sup>a)</sup>	0.064	0.244
missing human capital <sup>a)</sup>	0.215	0.411
manufacturing (technology intensive)	0.025	0.156
manufacturing (not technology intensive)	0.091	0.288
basic construction	0.086	0.280
construction (installation and completion)	0.096	0.294
car retailing and maintainance	0.054	0.225
retail trade	0.125	0.331
whole sale	0.221	0.415
transport & communication	0.062	0.240
business related services	0.126	0.332
consumption related services	0.114	0.318
ln(population density)	6.020	1.377
unemployment rate 1992	11.051	5.703
bank customer density in 1990/91	2.343	1.088
start-up in Western Germany	0.510	0.500
number of observations ( <i>N</i> )		10,827

*Note:* Federal States were dropped due to space restrictions, but are available on request.

<sup>a)</sup> Highest human capital endowment of the owner person(s)

very high - Professor or Doctor (PhD); high - first university degree,

master craftsman; medium - vocational training; low - unskilled worker;

missing - no details available in the ZEW Foundation Panels East and West

**Tabelle 2:** Descriptive Statistics for the observations used in the growth estimation

variable	mean/share	STDV	not assisted	assisted
number of employees at start-up	6.671	22.604	7.129	5.213
limited liability legal form	0.419	0.494	<b>0.467</b>	<b>0.294</b>
start-up in 1990	0.270	0.444	0.275	0.256
start-up in 1991	0.268	0.443	<b>0.249</b>	<b>0.318</b>
start-up in 1992	0.244	0.430	0.239	0.256
start-up in 1993	0.218	0.413	<b>0.237</b>	<b>0.169</b>
one firm owner	0.508	0.500	<b>0.493</b>	<b>0.547</b>
other firm(s) involved in start-up	0.088	0.283	<b>0.104</b>	<b>0.047</b>
very high human capital	0.037	0.188	0.041	0.025
high human capital	0.390	0.488	<b>0.367</b>	<b>0.451</b>
medium human capital	0.309	0.462	<b>0.326</b>	<b>0.265</b>
low human capital	0.041	0.197	0.045	0.030
missing human capital	0.224	0.417	0.222	0.229
manufacturing (techn. int.)	0.030	0.171	0.027	0.038
manufacturing (not techn. int.)	0.088	0.283	0.086	0.093
basic construction	0.074	0.262	<b>0.066</b>	<b>0.095</b>
installation and completion	0.119	0.324	<b>0.105</b>	<b>0.157</b>
car retailing and maintainance	0.053	0.225	0.049	0.066
retail trade	0.119	0.324	<b>0.135</b>	<b>0.076</b>
whole sale	0.232	0.422	<b>0.217</b>	<b>0.273</b>
transport & communication	0.060	0.238	0.063	0.053
business related services	0.158	0.364	<b>0.187</b>	<b>0.078</b>
consumption related services	0.066	0.248	0.065	0.070
ln(population density)	5.794	1.325	<b>5.960</b>	<b>5.355</b>
start-up in Western Germany	0.424	0.494	<b>0.514</b>	<b>0.186</b>
number of observations ( $N_1$ )	1,726		1,254	472

*Note:* bold numbers indicate statistical significance in a two-tailed t-test at the 5 per cent level, i.e. the means differ between both groups.

**Tabelle 3: Bivariate probit estimation**

variable	Receipt of assistance ( $D_i = 1$ )		Interview available ( $S_i = 1$ )	
	coefficient	standard error <sup>a)</sup>	coefficient	standard error <sup>a)</sup>
receipt of assistance ( $E[D_i X_{S_i}, X_{D_i}]$ )	/	/	0.774	0.301**
ln(number of employees at start-up)	0.207	0.034**	-0.119	0.043**
ln(number of employees at start-up) <sup>2</sup>	-0.045	0.008**	-0.017	0.011
limited liability legal form	-0.164	0.040**	0.026	0.042
start-up in 1991	0.141	0.039**	0.061	0.037
start-up in 1992	0.136	0.044**	0.053	0.046
start-up in 1993	-0.084	0.046	0.077	0.038*
one firm owner	0.061	0.031*	0.254	0.040**
other firm(s) involved in start-up	-0.453	0.060**	0.109	0.070
very high human capital	0.256	0.092**	0.116	0.090**
high human capital	0.357	0.036**	0.125	0.047
medium human capital	-0.064	0.090	-0.076	0.062
missing human capital	-0.052	0.045	0.098	0.042*
manufacturing (techn. int.)	0.160	0.103	0.165	0.097
manufacturing (not techn. int.)	0.054	0.054	0.053	0.062
basic construction	-0.032	0.062	-0.060	0.069
installation and completion	0.046	0.054	0.087	0.055
transport & communication	-0.144	0.070*	0.031	0.074
business related services	-0.266	0.057**	0.202	0.055**
consumption related services	-0.278	0.056**	-0.239	0.060**
ln(population density)	-0.088	0.018**	-0.047	0.013**
unemployment rate 1992	0.007	0.005	/	/
bank customer density in 1990/91	-0.227	0.096*	/	/
bank customer density <sup>2</sup>	0.030	0.016	/	/
start-up in Western Germany	/	/	-0.109	0.058
Schleswig-Holstein	-0.841	0.161**	/	/
Hamburg	-0.758	0.582	/	/
Lower Saxony	-0.877	0.102**	/	/
Bremen	-0.859	1.709	/	/
North Rhine-Westphalia	-0.766	0.073**	/	/
Hesse	-0.730	0.110**	/	/
Rhineland-Palatinate	-0.684	0.142**	/	/
Baden-Wuerttemberg	-0.675	0.109**	/	/
Bavaria	-0.781	0.094**	/	/
Saarland	-0.724	0.237**	/	/
Berlin	-0.396	0.129**	/	/
Brandenburg	-0.229	0.058**	/	/
Mecklenburg-Vorpommern	0.085	0.072	/	/
Saxony-Anhalt	-0.045	0.061	/	/
Thuringia	0.160	0.054**	/	/
intercept	0.018	0.169	-0.861	0.131**
correlation coefficient ( $\rho$ )			0.150**	
number of observations ( $N$ )			10,827	

Note: (\*) indicates statistical significance at the 1 (5) per cent level.

<sup>a)</sup> The standard errors were calculated using the bootstrap method.

**Tabelle 4:** Growth estimation to calculate the mean program impact

variable	coefficient	standard error	standard error (bootstrapped)
<b>mean program impact (<math>\hat{\theta}^1</math>)</b>	0.070	0.030*	0.029*
ln(number of employees at start-up)	-0.141	0.034**	0.038**
ln(number of employees at start-up) <sup>2</sup>	0.005	0.008	0.007
limited liability legal form	0.096	0.034**	0.035**
start-up in 1991	-0.005	0.034	0.033
start-up in 19992	-0.070	0.035*	0.026**
start-up in 19993	0.061	0.036	0.038
one firm owner	0.018	0.030	0.030
other firm(s) involved in start-up	0.028	0.048	0.048
very high human capital	0.196	0.073**	0.088*
high human capital	0.037	0.035	0.034
low human capital	-0.052	0.067	0.091
missing human capital	-0.047	0.036	0.033
manufacturing (techn. int.)	0.116	0.077	0.045*
manufacturing (not techn. int.)	0.197	0.047**	0.065**
main construction	0.159	0.053**	0.063*
installation and completion	0.158	0.043**	0.032**
transport & communication	0.102	0.054	0.056
business related services	0.178	0.041**	0.055**
consumption related services	-0.044	0.059	0.062
ln(population density)	-0.021	0.011	0.010*
start-up in Western Germany	-0.104	0.032**	0.033**
Mill's ratio ( $\hat{\lambda}_{Si}$ )	0.252	0.099*	0.096**
intercept	-0.094	0.129	0.111
adjusted $R^2$		0.056	
number of observations ( $N_1$ )		1,726	

*Note:* \*\*(\*) indicates statistical significance at the 1 (5) per cent level.