Are efficiency measures predictive of firm crisis? Evidence from the Italian agri-food industry

1. Introduction

Firm survival is one of the main issues in the field of research related to entrepreneurship, organizations, and business management resulting in a large and heterogeneous body of scholarly literature (Audretsch and Mahmood 1995; Josefy et al. 2017; Suarez and Utterback, 1995).

In the last decades scientific efforts concentrated on the determinants related to the survival of firms. Earlier research has faced this issue by arguing that the success and survival are potentially driven by very different conditions (e.g. Cooper 1993; Kalleberg and Leicht 1991).

The present study represents an advance in the firm survival research field (e.g., Audretsch and Mahmood 1995; Geroski et al. 2010; Josefy et al. 2017; Strotmann 2007) by examining the role of efficiency and productivity in the context of surviving attitude of firms in specific industries, and testing if it could be considered as a predictive symptom.

The agri-food industry, particularly in Italy, is an example of an industry that has experienced a notable growth during the last years in terms of economic significance, number of dedicated firms and geographical scope of activities. However, commonly due to their small size, the main goal, for a large number of agri-food firms is to ensure their survival.

If agri-food firms seek to ensure their survival and Italian public institutions wish to support a vivid agri-food industry, identifying the conditions that specifically help firm survival is crucial. The paper explore the effects of some scientifically established determinants on Italian agri-food firm survival, adding another factor that is the production efficiency, testing if the latter affect or not the probability of firm default. Our results provide important policy relevant knowledge of what institutional, industry- and firm-level factors should be developed in order to keep the firms alive.

2. Literature review on the survival analysis of firms

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Title</th>
<th>Journal</th>
<th>Model</th>
<th>Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oakes (1982)</td>
<td><strong>Survival Analysis</strong></td>
<td>European Journal of operational research</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late et al. (1986)</td>
<td><strong>An application of the cox proportional hazards model to bank failure</strong></td>
<td>Journal of Banking and Finance</td>
<td>cox and MDA</td>
<td>Capital, asset quality, loan composition, efficiency (pricing), earnings, liquidity (see pag 517)</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Title</td>
<td>Journal/Book</td>
<td>Method(s)</td>
<td>Variables</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Luoma Laitinen (1993)</td>
<td>Survival Analysis as a tool for company failure prediction</td>
<td>Omega</td>
<td>cox</td>
<td>ROI, CFS(chash flow to net sales ratio), APA (Average payment period for accounts payable), SCT(Shareholders capital to total capital ratio), OEC(Official expenses to corrected expenses ratio), LNS(logarithmic net sales)</td>
</tr>
<tr>
<td>Laitinen (1993)</td>
<td>Financial predictors for different phases of the failure process</td>
<td>Omega</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartelsman et al. (2005)</td>
<td>Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries</td>
<td>Industrial and corporate change</td>
<td>fixed effect regression</td>
<td>countries</td>
</tr>
<tr>
<td>Authors</td>
<td>Title</td>
<td>Journal</td>
<td>Model(s)</td>
<td>Predictors</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Holmes et al. (2008)</td>
<td>Analysis of new firm survival using a hazard function</td>
<td>Applied Economics</td>
<td>log logistic hazard model</td>
<td>plant size, real interest rate, change in real interest rate (1/2/3 years), sectoral growth, manufacturing code</td>
</tr>
<tr>
<td>Daep et al. (2015)</td>
<td>The mortality of companies</td>
<td>Interface</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audretsch et al. (2016)</td>
<td>Ownership, productivity and firm survival in China</td>
<td>Economia e politica industriale</td>
<td>cox</td>
<td>ownership, soft budget constraint, productivity, firm size, innovation, export, industry growth and stage of business cycle.</td>
</tr>
<tr>
<td>Arbia et al. (2016)</td>
<td>A spatial analysis of health and pharmaceutical firm survival</td>
<td>Journal of applied statistics</td>
<td>cloglog model, exponential model, cox, complementary log log, complementary log log mixed</td>
<td>dispersion, independence, concentration, employees, legal status, region, log incidence rate</td>
</tr>
<tr>
<td>Hakanson &amp; Kappen (2016)</td>
<td>Live and let die: a survival analysis of foreign R&amp;D units in Swedish MNEs</td>
<td>International business review</td>
<td>cox</td>
<td>establishment mode, integration, local embeddedness, global mandate, autonomy, size, age, local market growth, country is english speaking, country IPR</td>
</tr>
</tbody>
</table>
### 3. Research design

The study aims at analyzing the agrifood industry in Italy for the period 2009/2018. According to the established scientific literature this time period is sufficiently long to highlight the main factors affecting firm survival. It includes the years of the global financial crisis, which had negative impacts on each Italian economic sector. Although the effects of the crisis could not be accurately controlled and excluded from the analysis, the selected time period can capture the firm and industry-specific determinants of survival. The following picture is the conceptual framework scheme:

The dataset collect 355 firms operating in the agrifood sector and is constructed by extracting data from AIDA Bureau Van Dijk database, selecting the period of analysis and the activity code manufacture of agri-food products in line with Nace Rev.2, adopted by a large scientific literature. It includes firm exit dates, and longitudinal data on firm characteristics such as size, ownership status, and location, extracted from the above-mentioned sources.

The sample is constituted by the total of firms dissolved during the period 2009/2017, in Italy, belonging to sector 10 of Nace Rev. 2 and non-dissolved firms, identified through the propensity score matching on the basis of the following observable characteristics:
- revenues in 2009 (first year of the analysis);
- province;
- activity code.

Non-dissolved firms have been considered right-censored.

### 3.1. Data Envelopment methods and conditional efficiency model

Following Cazals et al. (2002) and Daraio and Simar (2005) we consider a production technology where the activity of the production units is characterized by a set of inputs $X \in \mathbb{R}_+^p$ used to produce a set of outputs $Y \in \mathbb{R}_+^q$ and $Z \in \mathbb{R}_+^d$, generic vector of environmental variables.
Mastromarco and Simar (2015) extend this framework to the time dimension. We follow their approach and consider a panel of data \((x_i, t; y_i, t; z_i, t)\) for \(i = 1, \ldots, n\) and \(t = 1, \ldots, s\) the unconditional and conditional attainable sets can be estimated. To introduce the time dimension we will consider the time \(T\) as an additional conditioning variable and, for each time period \(t\), define the attainable set \(\Psi^T_t \subset \mathbb{R}^{p+q}_+\) as the support of the conditional probability:

\[
H^T_{X,Y|Z}(x,y|z) = \text{Prob}(X \leq x, Y \geq y | Z = z, T = t)
\]

Accordingly, the conditional output-oriented technical efficiency of a production plan \((x; y) \in \Psi^T_t\), at time \(t\) facing conditions \(z\), is defined in (Daraio and Simar, 2005) as:

\[
\lambda_t(x,y|z) = \sup \{\lambda(x,\lambda y) \in \Psi^T_t \} = \sup \{\lambda | S^T_{y|x,z}(\lambda y|x,z) > 0\}
\]

where \(S^T_{X,Y|Z}(x,y|z) = \text{Prob}(Y \geq y | X \leq x, Z = z, T = t)\).

Assuming that the true attainable sets are convex and under free disposability of inputs and outputs, the DEA estimators can be written as (Daraio and Simar, 2007b)

\[
\Psi_{DEA} = \left\{ (x,y) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+ \mid \begin{array}{l}
y \leq \sum_{j=1}^n y_j x_j \\
\sum_{j=1}^n y_j = 1
\end{array} \right\}
\]

where the summations are over all the combinations \(j = (i; t)\) with \(i = 1, \ldots, n\) and \(t = 1, \ldots, s\). Similarly, at time \(t\) and facing the conditions \(Z = z\), we have

\[
\Psi^T_{DEA} = \left\{ (x,y) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+ \mid \begin{array}{l}
y \leq \sum_{j=1}^n y_j x_j \\
\sum_{j=1}^n y_j = 1
\end{array} \right\}
\]

where \(J(z; t) = \{j = (i; v) \mid z - hz < z_i, v < z + hz; t - ht < v < t + ht\}\); \(hz\) and \(ht\) are bandwidths of appropriate size selected by data-driven methods. These \(J(z; t)\) describe the localizing procedure to estimate the conditional DEA estimates and they determine the data points in a neighbourhood of \((z; t)\) that will be used to compute the local DEA estimate. The choice of appropriate bandwidth selection is the focus of Daraio and Simar (2005, 2007a) and Badin et al (2010) studies. They are determined by the estimation of conditional distributions \(S^T_{y|x,z}(y|x,z)\) conditioned on \(X \leq x\), time \(T = t\) and a particular value of \(Z = z\) adapting standard tools from Hall et al. (2004) and Li and Racine (2007). Of course, here only the variables \((t; z)\) require smoothing and appropriate bandwidths, since we have

\[
\tilde{S}^T_{y|x,z}(y|x,z) = \frac{\sum_{j=1}^n I(x_j \leq x, y_j \geq y) K_{hz}(z_j - z) K_{ht}(v - t)}{\sum_{j=1}^n I(x_j \leq x) K_{hz}(z_j - z) K_{ht}(v - t)}
\]

where the function are kernels with compact support (see Badin et al., 2010, for technical details). Optimal bandwidths can be selected by least squares cross-validation (LSCV) or by maximum likelihood cross-validation, which are asymptotically equivalent (see, for example, Li and Racine (2007)).

The effects of conditioning variables on the boundary and on the distribution of the inefficiencies can be identified using the approach developed by Badin et al. (2012). The effect on the boundary can be detected by assessing the ratios between conditional to unconditional efficiency measures, relative to the full frontier of the conditional and the unconditional attainable sets:

\[
R^T_0(x,y|z, t) = \frac{\lambda_t(x,y|z)}{\lambda(x,y)}
\]

The focus of our study are, in particular, the effects of \(T\) and \(Z\) on these ratios. For the output orientation, the conditional efficient boundary is by construction below the unconditional one, while ratio is equal to 1, if and only if there is no shift of the efficient boundary of the two attainable sets, at time \(t\) and with conditions \(z\). Examining the potential differences between the boundaries of the attainable sets and assessing the above ratio, we can just establish if the environmental variables affect the technology: if it is increasing, it indicates positive effect. For the evaluation of the influence of the distribution of the inefficiencies below the frontier, we have to use partial frontiers.
We may also look, as suggested in Badin et al. (2012) to the order-\(\alpha\) counterparts by choosing more central quantiles, such as the median, to investigate the effect of \(T\) and \(Z\) on the distribution of inefficiencies. The ratios to be analyzed are now

\[
R_{0,\alpha}(x, y | z, t) = \frac{\lambda_{t,\alpha}(x, y | z)}{\lambda_{\alpha}(x, y)}
\]

Some potential shifting effect already observed with equation (first) could be enhanced (or reduced) if the effect is different with the ratios (second) computed for smaller \(\alpha\).

For smaller \(\alpha\), if the effect is different, some potential shifting effect could be enhanced (or reduced). Since the order-\(\alpha\) efficiency scores are not bounded by 1, also the ratios are not bounded by 1 (Badin et al., 2012).

In practice, we use nonparametric estimators of the efficiency scores and we explore the effect of \(T\) and \(Z\) by looking at the behaviour of \(R_{0}(x, y | z, t)\) and \(R_{0,\alpha}(x, y | z, t)\) as a function of \(T\) and \(Z\).

To analyze the effect of our external variables we use a non-parametric model for the second stage regression. However, this two-stage approach requires a restrictive separability condition between the input/output space and the space of external factors. In the absence of separability condition, the marginal measures of efficiencies are economically meaningless because each observed production plan is compared with an unattainable frontier at any particular time \(t\) and facing conditions \(z\) (Simar and Wilson, 2007; 2011). In this case, the regression of the conditional efficiency scores on the relevant explanatory variables in the second stage is more meaningful (Badin et al., 2012).

The shift of the frontier and technological changes, caused by time and values of \(Z\), could be captured in the efficiency scores.

As explained in Simar and Wilson (2007; 2011) the “separability” assumption assumes that the frontier of the attainable set does not depend on the values of \(Z\).

Simar and Wilson (2007) note (pp. 35-36) that their Assumptions A1-A2 imply a “separability condition”, and this condition may or may not be supported by the data, and hence that the condition should be tested.

Daraio et al. (2018) provide a fully non-parametric test of this condition. If it is rejected, then the conditional efficiency measures as explained by Daraio and Simar (2005, 2006) are appropriate. Daraio et al. (2018) by developing central limit theory (CLT) and, using the CLT results for both unconditional and conditional efficiencies, suggest a test of separability assumption \(\tau(x; y) = \tau t (x; y | z)\) versus the alternative of non-separability. They compare the conditional and unconditional efficiency scores using relevant statistics of \(\tau(x; y)\) and \(\tau t (x; y | z)\). Following Daraio et al. (2017) to implement the test we randomly split our data sample in two independent parts \(n_1\) and \(n_2\) and use them to compute conditional and unconditional estimates. Then we compute the estimates of the average and variances in the two independent samples:

\[
\hat{\mu}_{n_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} \tilde{\tau}_i (x; y)
\]

\[
\hat{\mu}_{n_2} = \frac{1}{n_2} \sum_{i=1}^{n_2} \tilde{\tau}_i (x; y | z)
\]

\[
\hat{\sigma}^2_{n_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} (\tilde{\tau}_i (x; y) - \hat{\mu}_{n_1})^2
\]

\[
\hat{\sigma}^2_{n_2} = \frac{1}{n_2} \sum_{i=1}^{n_2} (\tilde{\tau}_i (x; y | z) - \hat{\mu}_{n_2})^2
\]

Then we estimates also the bias for a split of each subsample for the unconditional \(\hat{\beta}_{n_1}\) and conditional cases \(\hat{\beta}_{n_2}\). Then we calculate the asymptotically normal test statistic:

\[
Z = \frac{(\hat{\mu}_{n_1} - \hat{\mu}_{n_2}) - (\hat{\beta}_{n_1} - \hat{\beta}_{n_2})}{\sqrt{\frac{\hat{\sigma}^2_{n_1}}{n_1} + \frac{\hat{\sigma}^2_{n_2}}{n_2}}} \xrightarrow{\text{N} (0; 1)} \frac{1}{\sqrt{\frac{n_1}{n_1} + \frac{n_2}{n_2}}}
\]
Under the null, because the bias-corrected sample means are independent and two sequences of independent normal limiting distributed variables have a joint bivariate normal distribution, the statistic is distributed as a standard normal. We thus regress, in the second stage, the conditional efficiency scores on size and time. This is motivated by the fact that the so-called “separability condition” discussed in Simar and Wilson (2007) likely does not hold.

3.1.1. Data collection, variables construction and model implementation

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>Z variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of employees</td>
<td>Added value</td>
<td>Time</td>
</tr>
<tr>
<td>Material Asset</td>
<td></td>
<td>Indicator of infrastructure at province level</td>
</tr>
</tbody>
</table>

The DEA conditional efficiency approach considers, besides the traditional input and output of the production function, also two environmental variable, i.e. the time and the infrastructural endowment index. The level of competitiveness and attractiveness of a territory is determined by the adequacy of the economic and social infrastructures present in the reference areas. The presence of infrastructures represents a relevant contextual factors for firms operating in the reference territory, showing, indeed, a positive correlation with per capita income levels.

Infrastructures are a heterogeneous category and concern:
- road network;
- railway network;
- ports;
- airports;
- energy and environmental systems and networks;
- telephony and telematics structures and networks;
- banking networks and business services;
- cultural and recreational facilities;
- educational facilities;
- health facilities.

The procedure followed for the construction of the synthetic index of infrastructure endowment is as follows. After transforming the elementary variables into quotas on the national total, two indicators of “absorption”, one of quantity and one of quality, were built through a weighted average of the elementary variables. In order to identify the weight system in a non-discretionary way, a principal component analysis (PCA) was carried out using the correlation matrix. The identified weights are proportional to the factor coefficients of the first component.

The overall qualitative and quantitative infrastructural endowment index at the provincial level was then determined by the weighted arithmetic mean of the two indicators with inversely proportional weights to a measure of their variability. In this way, a greater weight has been implicitly assigned to the quantitative indicator, which usually has less variability. It is noteworthy that, having carried out the calculations on the elementary series expressed in percentages, also the synthetic index of infrastructural endowment is represented by a series of data that indicate the quota on the national total pertaining to each province.

In order to neutralize the effects due to the different territorial dimension, the endowment indicator must be related to a similar indicator of potential demand, expressed by the population, by the surface or an actual demand indicator. The ratio between supply indicator and demand indicator determines territorially comparable comparative endowment indices and provides a value equal to 100 for the entire national economy and, respectively, values higher or lower than 100 depending on whether they are territories with a relative endowment higher or lower than the national average. (source: https://www.ucer.camcom.it/studi-ricerche/dati/bd/infrastr/numeri-indici-delle-dotazioni-infrastrutturali-n-r-p)
3.2. Survival analysis methodology

Denote \( T \) a non-negative continuous random variable, representing the time until the event of interest. The distribution function is \( F(t) = P(T \leq t) \) while \( f(t) \) is the probability density function. In order to conduct a survival analysis we have to know the survival function \( S(t) \), the cumulative hazard function \( H(t) \), the hazard function \( h(t) \) and the mean residual life function \( mrl(t) \). It is sufficient to know one of these in order to determine the others. The survival function \( S(t) \) represents the probability that a randomly selected individual will survive beyond time \( t \) and is defined as follows:

\[
S(t) = P(T > t) = 1 - F(t)
\]

It is a decreasing function which has a codomain in \([0; 1]\), in particular assuming value equal to 1 at \( t = 0 \) and equal to 0 at \( t = 1 \). The cumulative hazard function \( H(t) \) is an increasing function, assuming values in \([0; +\infty]\) defined as follows:

\[
S(t) = \exp(-H(t))
\]

The hazard function is a positive function and measures the instantaneous risk of exiting from the market right after time \( t \) given that the individual is alive at time \( t \). It can have many different shapes. It is therefore a useful tool to summarize survival data and can be defined as follows:

\[
h(t) = \frac{d}{dt}H(t)
\]

Since often time-to-event data are incomplete, different kind of analysis can be conducted. The nature of our study suggests that we are in the presence of a random right censoring. In fact, in our case censoring is a random variable and the study itself continues until a fixed point in time, but firms enter and leave the study at different times. Since classical methods do not work for censored data, such as that in our case study, nonparametric estimators of the survival function can be adopted, i.e. cumulative hazard function or hazard rate (a measure of association). These estimators are based on the nonparametric likelihood function that is different from the likelihood for completely observed data due to the presence of a censoring indicator.

Given a random sample of individuals of size \( n \), the survival time is denoted with \( T_1, \ldots, T_n \) while the censoring time with \( C_1, \ldots, C_n \). Given observed data \((Y_i, \delta_i) (i = 1, \ldots, n)\) with \( Y_i = \min(T_i, C_i) \) and \( \delta_i = I(T_i < C_i) \) let us denote \( f(\cdot) \) and \( F(\cdot) \) for the density and distribution of \( C \), assuming that \( T \) and \( C \) are independent, the formula of the likelihood is:

\[
\prod_{i=1}^{n} [(1 - G(Y_i))f(Y_i)]^{\delta_i} [(1 - F(Y_i))g(Y_i)]^{1-\delta_i}
\]

or it can also be written as follows:

\[
L = \prod_{i \in D} f(y_i) \prod_{i \in R} S(y_i)
\]

where \( D \) represents the index set of survival times and \( R \) represents the index set of right censored times.

Survival models consist of the underlying hazard function, often denoted \( h_0(t) \), describing how the risk of event per time unit changes over time at baseline levels of covariates and the effect parameters, describing how the hazard varies in response to explanatory covariates. Cox (1972; 1975) observed that if the proportional hazards assumption holds (or, is assumed to hold) then it is possible to estimate the effect parameter(s) without any consideration of the hazard function. This approach to survival data is called Cox proportional hazards model, sometimes abbreviated as Cox model or proportional hazards model.

In particular, this study employed the semi-parametric Cox proportional hazards model (Cox 1972; 1975) to estimate the models for the dependent variable.

The Cox model specifies the hazard of firm exit as:

\[
h(t, x(t), \beta) = h_0(t) e^{x'(t)\beta}
\]
where $h_0(t)$ is the unknown baseline hazard function, $x(t)$ denotes the vector of covariates expected to shift the hazard of exit proportionally in each year, and $\beta$ is a vector of parameters to be estimated (Hosmer and Lemeshow 1999). The advantage of using the Cox proportional hazards model is that one does not need to make parametric assumptions about the form of duration dependence in the hazard rate (Cleves et al. 2004). This analytic method has been used by Audretsch and Mahmood (1995) to analyze new firm survival and, for example, by Boyer and Blazy (2014) in their analysis of survival of micro-startups. In addition, the findings of Cader and Leatherman (2011) provide strong support for the use of Cox hazard models for firm survival analysis.

### 3.2.1. Data collection, covariates and model implementation

In the second phase of this study, a COX proportional Hazard Regression is conducted on the statistical units, i.e. firms included in the dataset. Some firm-specific features, such as the location, the size, the results of the conditional efficiency analysis (conditional efficiency score, unconditional efficiency score and managerial (or pure) efficiency score) and the Altman score, are identified as covariates in the survival analysis, while the duration, measured in year, as dependent variable. Since survival analysis does not include the time variable, covariates (such as Altman scores and DEA scores), need to be reduced to one firm-specific and time-invariant value for example through the geometric mean.

The dependent variable is directly measured by the firm survival, computed on the basis of the firm exit date from the market, due to its inability to survive (e.g., Agarwal and Sarkar 2002; Klepper 2002). The study defined exit from the market, the failure of a firm (cf. Baum 1996), i.e. bankrupt or value added tax liability expiry. As already established in the specific-field literature, mergers and acquisition, implying discontinuance of a firm have been treated as right-censored cases.

To conduct a survival analysis on Italian agri-food firms, we used variables which are proved to be determinants of firm crisis according to the established business literature and in particular to Altman model, adjusted to capture SME’s peculiarities (Altman, 1111). The predictive model of Altman known as the Z-Score test allows to predict, with statistical techniques, the probability of failure of a company in the following years, deriving the data from the financial statements and based on the discriminating statistical analysis. The covariates used in the survival analysis of the present study include the variables considered by Altman for the Z-score test and in addition the measure of the conditional efficiency calculated using the DEA methodology on the same sample of companies.

The first covariate, indicated with $X_1$, expresses the value of the firm’s liquid assets with respect to total capitalization. It is clear that a firm that faces substantial operating losses will have a sharp reduction in current assets in relation to total assets. This index proved to be the best among the liquidity indices tested, including the current ratio and the quick ratio.

\[
X_1 = (C_A - C_L) / (M_A + I_A + F_I + C_A + C_E)
\]

The covariate named $X_2$ expresses the firm ability to reinvest its profits. A young firm will certainly have a lower index compared to a more ancient one; This is because the young firm has not yet had time to build up reserves, and therefore may be penalized in assessing the risk of bankruptcy. This is precisely the real situation in which the newly constituted firms have a greater probability of failure in the first years of their life.

\[
X_2 = (L_R + E_R)/T_A
\]

The $X_3$ covariate measures the productivity of the activities of a firm, without any financial or fiscal leverage factor. For this reason, this index is particularly appropriate in defining the probability of insolvency.

\[
X_3 = N_{O1}/(M_A + I_A + F_I + C_A)
\]

The covariate indicted with $X_4$ shows how much the activities of a firm can be reduced before the total liabilities exceed the assets and the conditions for bankruptcy are created.

\[
X_4 = E/T_L
\]
The last factor of Altman model, named covariate $X_5$ in our study, measures the ability of a firm to generate revenues given a certain asset value. It expresses the entrepreneurial capacity to relate to the competitiveness of the firm’s reference market.

$$X_5 = \frac{S_R}{(M_A + I_A + F_i + AC + DL)}$$

Among the above-mentioned covariates, this study identifies also a measure of efficiency in terms of firm productivity. The addition of this further variable allows this study to be considered original and novel. In fact the main aim of this paper is to test if efficiency measures are predictive of firm crisis. The productivity variable consists of a score deriving from the conditional efficiency analysis conducted through the DEA approach using as Decision making Units, the same units of sample considered for the following survival analysis.