

An Analysis of Factors Causing Farms Located in the Veneto Region to Stop Productive Activity

Un'Analisi dei Fattori di Rischio di Cessare l'Attività Produttiva per le Aziende Agricole della Regione Veneto.

Francesca Bassi*, Ottorino Chillemi**, Adriano Paggiaro*

*Statistics Department, University of Padova, via C. Battisti 241, 35121, Padova, Italy;
francesca.bassi@unipd.it, adriano.paggiaro@unipd.it

**Economics Department, University of Padova, via del Santo 22, 35123 Padova, Italy,
ottorino.chillemi@unipd.it

Riassunto: in questo lavoro si analizzano i fattori che favoriscono la sopravvivenza delle aziende agricole venete allo scopo di prevederne il rischio di cessare l'attività produttiva. I dati utilizzati provengono da una nuova base di dati che abbina le informazioni contenute nel registro imprese delle Camere di Commercio con quelle raccolte durante il Censimento dell'Agricoltura del 2000.

Keywords: survival analysis, subsidies to agriculture, scoring functions

In the immediate future, State aids for the agricultural sector are expected to be granted on a more selective basis. This makes it interesting to measure farms' survival ability. (Consiglio UE, 2006; Regione Veneto, 2006). In this paper, we consider a farm to be vital if it survives for at least 5 years, relaying on the widely shared opinion that it is not advisable to subsidize economic activities which are predicted to close within the following 5 years (Key and Roberts, 2006).

Our analysis uses data from a new database which links information on farms' survival collected in the Official Registers of the Chamber of Commerce (REA archives) with that collected in the Agricultural Census on socio-economic characteristics of farms (Istat, 2000). The database contains information on 91.244 farms operating in the Veneto region (North-east of Italy) which are present in the REA archives in 1999 and were interviewed during the 2000 Census. From the REA archives farms' survival history is followed, on a yearly basis, till 2004.

After eliminating few records with missing data, the remaining farms (90.702) were randomly divided into two equivalent samples (45.351 units each), in order to check forecasting performance. We estimated on half sample a logit model for the event of stopping productive activity during a period of 5 years after 1999, using as covariates some socio-economic characteristics of farms collected in the Census and age of the farm recoded in classes. The estimated probability of closing should be used as a score, and a threshold has to be chosen in order to decide which farms are to be financed.

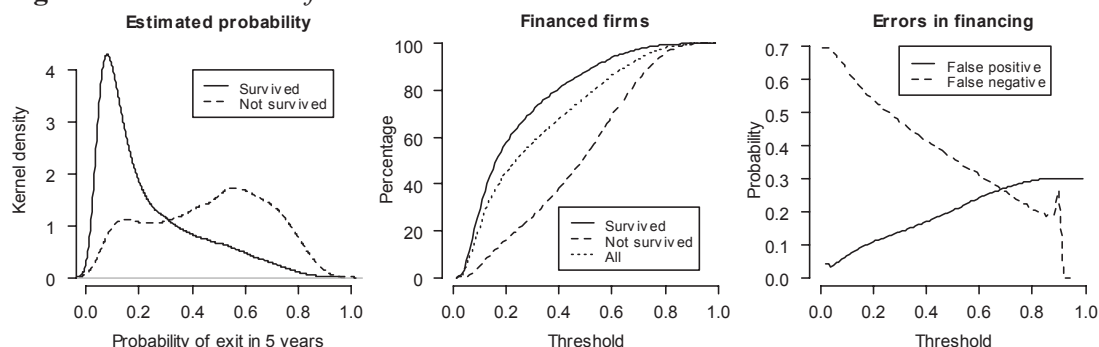
Before presenting estimation results, two comments are in order: (i) prediction on the two samples gives almost the same results; (ii) estimation of a discrete time duration model, with and without unobserved heterogeneity (Jenkins, 1995), obtains very similar results. Thus, we present the performance of the simple logit model on the test sample.

Many of the covariates in the model turn out to have significant effects on the probability of closing. The most effective are those related to size, sector and province, with bigger farms surviving longer than smaller ones. Finally, once we control for the

other characteristics farm age is no more significant, and this partly explains why duration models do not improve predictions.

Figure 1 contains, for both survived and not survived farms, the kernel density of the estimated probability of closing and the percentage of subsidized firms for each given threshold; moreover, it presents the rate of false positives (financed farms not surviving after 5 years) and false negatives (not financed farms surviving after 5 years).

Figure 1: Main results of model estimation



The effectiveness of the scoring method to control for classification errors depends on the threshold: if we finance all farms (threshold 1), the false positive rate is at its maximum .30 (the overall observed transition rate), while if we finance no farm (threshold 0), the false negative rate is at its maximum $1 - .30 = .70$.

There are many possible strategies in order to choose the most appropriate threshold, depending on which cost function is associated to the decision. Among these, Table 1 shows some examples of different simple strategies:

- maximum number of farms which can be financed (60% in the example);
- maximum acceptable rate of false positives (10%);
- maximum acceptable rate of false negatives (30%).

Table 1: Performance of the model under different strategies for choosing the threshold

Strategy	Thresh.	% Financed firms			Errors in financing	
		All	Survived	Not surv.	False pos.	False neg.
Financed 60%	0.32	59.98	73.60	28.41	14.28	46.09
Max f.p. 10%	0.18	41.51	53.49	13.74	9.98	55.55
Max f.n. 30%	0.62	88.14	94.92	72.45	24.77	29.96

References

- Consiglio UE (2006) *Decisione del Consiglio – 20/02/2006*, G.U. Unione Europea.
- Istat (2000) *Piano generale del V Censimento dell'Agricoltura*.
- Jenkins S. P. (1995) Easy ways to estimate discrete time duration models, *Oxford Bulletin of Economics and Statistics*, 57, 129-138.
- Key N., Roberts M. J. (2006) Government payments and farm business survival, *American Journal of Agricultural Economics*, 88, 382-392.
- Regione Veneto (2006) *Dgr. n. 1189, 02/05/2006 Programmazione e sviluppo rurale 2007-2013*.