

DETERMINANTS OF THE DURATION OF CONTRACTS: A LONGITUDINAL STUDY IN THE ITALIAN FISHERY SECTOR

Luciano Nieddu, Cecilia Vitiello

1. Introduction

In Italy the average age of workers with non-permanent position is increasing: the transition from a non-standard contract to a permanent position is more and more unlikely. In the past decades a non-standard contract was often a first step to gain a permanent job position. This is rapidly changing and the non-standard contract is more likely to become an absorbing state. In this framework the aim of this paper is to study the duration of contracts in the Italian fishery sector in the last decade in order to determine whether the previous work history of an employee influences the duration of her/his future contracts.

We will focus our attention on administrative data from the National Institute of Social Insurance (INPS). A stratified sample of contracts from the Marche region has been analyzed. For each contract the length and the previous work history of the corresponding employee is available (longitudinal data).

The aim of the paper is twofold: on one side we face the methodological issue of summarizing the previous work history of an employee, on the other we try to determine if the previous work history influences the duration of future contracts.

The structure of the paper is as follows: in Section 2 the data used in the analysis will be described. In Section 3 the Cox model and its relevant extensions will be briefly recalled while in Section 4 the results will be presented. Finally in Section 5 some conclusions will be drawn.

2. The data

The data considered in this paper are administrative data collected from 2000 to 2009 and refer to all the work contracts that have been issued in the fishery sector. Our attention will be focused on the Marche region which represents the third most important region for the fishery sector in Italy. Given the administrative nature of the data, it must be stressed that they only describe the official situation for the fishery sector, depicting a reality which is the effect of policy maker decisions,

legislative choices and labor contract agreements. The statistical unit for the data set considered is the work contract. For each statistical unit the available information is: year, duration of the contract, age at the beginning of the contract, gender, work position, economic sector, firm issuing the contract.

The data set is composed of 16135 contracts; only 12 referring to employees with a position of manager or executive and only 14 from the economic sector known as "fish farming". In order to control for covariates such gender, economic sector and position, a stratified sample of contracts has been considered, dropping contracts that refer to executives or managers or associated to workers in the fish-farming sector. Once a contract was selected for the study, all the history of the worker in terms of previous and future contracts was selected to be part of the study. In Table 1 the sample is described with respect to gender, economic sector and position.

Table 1 –Sample structure: controlling for Gender, Economic Sector and Position.

POSITION	MALES				FEMALES			
	SECTOR				SECTOR			
	Commerce	Curing	ISF*	TOT	Commerce	Curing	ISF*	TOT.
Apprentices.	33	36	2	71	33	44	1	78
White Collars(WC)	40	39	83	162	51	56	16	123
Blue Collars(BC)	55	44	83	182	45	49	91	185
TOTAL	128	119	168	415	129	149	108	386

Note: (*) ISF: Internal and Sea Fishery

The marginal distributions in Table 1 seem to be fairly balanced, although the number of sample units for some configurations of the factors could not be balanced due to very small number of units in the corresponding population.

The 801 (415+386) contracts refer to 462 workers that had at least a contract in the period from 2000 to 2009. The record for each worker is therefore a string of length 120 indicating, for each month, if the worker has been employed or not.

At the beginning of each contract the history of the employee can be summarized in various ways: for instance considering indexes that ignore the order in which employment and unemployment periods occur, i.e computing average duration of contracts, duration of unemployment, average number of contracts prior to the current contract, etc. Another possibility is to consider a synthesis of the previous history of an employee, trying to summarize the pattern of his previous contracts, and indirectly assuming that employees with the same past work pattern are likely to experience the same pattern in the future.

The proposed methodology consists in determining the pattern in the history of a worker and to group workers with similar patterns. Contract histories are of

different lengths and of different start and endpoints, making difficult a direct comparison. Therefore various synthesis of the pattern have been studied. For instance, consider the data in Table 2 which refers to a 15-month period in the history of a person. The worker considered in the example has obtained a contract at time t13 and has been unemployed for 3 months prior to this contract. In the second row of the table, the cumulative number of employment months has been computed (H) while, in the third row, the cumulative number of months - regardless of employment or unemployment - from the starting period has been displayed (R). In the fourth and fifth row the ratio between the cumulative number of worked months and, respectively, the cumulative number of months and the total number of months have been displayed (“da” and “dac” respectively).

Table 2 –Summarizing previous history: an example for a contract starting at time 13.

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12	t13	t14	t15
t	0	1	1	1	0	0	1	1	1	0	0	0	1	1	1
H	0	1	2	3	3	3	4	5	6	6	6	6	-	-	-
R	1	2	3	4	5	6	7	8	9	10	11	12	-	-	-
da	0.0	0.5	0.7	0.8	0.6	0.5	0.6	0.6	0.7	0.6	0.5	0.5	-	-	-
dac	0	0.1	0.2	0.2	0.2	0.2	0.3	0.4	0.5	0.5	0.5	0.5	-	-	-

More formally, consider a contract starting at time T^* and let

$$t_{i,j} = \begin{cases} 0 & \text{if worker } i \text{ has not worked at time } j \\ 1 & \text{if worker } i \text{ has worked at time } j \end{cases}$$

$$da_{i,j} = \sum_{r=1}^j \frac{t_{i,r}}{r}; \quad dac_{i,j} = \sum_{r=1}^j \frac{t_{i,r}}{T^*}; \quad i = 1, \dots, n; j = 1, \dots, T^* \quad (1)$$

where $da_{i,j}$ compares a worker with a work history with no interruption and $dac_{i,j}$ is the empirical cumulative density function.

For each contract starting at time T^* the averages and the standard deviations of the two distributions can be computed, resulting in four indicators describing the work history prior to the current contract. The averages of $da_{i,j}$ and $dac_{i,j}$ are highly correlated so only the former will be considered in the study. Such indicators cannot be evaluated for contracts starting prior to 2000.

3. Survival Analysis

In survival analysis dependence of survival time from covariates is expressed modeling the hazard function $\lambda(t, x)$:

$$\lambda(t, x) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T < t + \Delta t)}{\Delta t} \quad (2)$$

The Cox proportional hazard model assumes that covariates are multiplicatively related to the hazard and that the hazard function can be factories as $\lambda(t, x) = \lambda_0(t)e^{x\beta}$, where $\lambda_0(t)$ is a function of time and $x\beta$ is a linear predictor not depending on time.

We need to consider extensions of the Cox model allowing for lack of proportionality and lack of independence. The information regarding the order in which contracts occur for an individual cannot be used, in fact while it has a precise meaning within a single work history it is not comparable with that of another individual because only a time windows of 10 years is available for each subject and not his whole history. Therefore the straightforward choice to treat dependence is via a frailty $\lambda(t, x) = \lambda_0(t)e^{x_j\beta + w_j}$.

In case on non-proportionality one possible solution is to divide the subjects into disjoint strata, each having a distinct baseline hazard function but common values for the coefficient vector β . The hazard for an individual j belonging to stratum k is $\lambda_k(t, x) = \lambda_{0k}(t)e^{x_j\beta + w_j}$.

The final outcome of the model is therefore a set of baseline hazard functions, one for each of the combination of the stratifying variables and a common estimate for the coefficient vector β . No probabilistic assertion can be made on the effect of stratifying covariates.

4. Results

As it is the case in many survival analysis problems, the data set considered in this paper presents left truncated observations. This has been taken into account in the model introducing a dummy variable (“old”) equal to 1 if the contract was already existing at 1st January 2000.

We have first considered a Cox model expressing the duration of contracts as a function of the structural variables in the dataset and some information on the previous history (“age at the beginning of the contract”-“age start” and “if the worker had been unemployed the month prior to the new contract-D1”). The proportionality test based on the scaled Schoenfeld residuals for the whole model (see Table 3) leads to a rejection of proportionality, gender being the factor with the lowest p-value.

Therefore a stratified Cox model, with gender as stratum, has been considered. The proportionality test for the new model leads again to a rejection of the null hypothesis (see Table 3). The covariate with the lowest p-value (D1) has been

selected as a candidate stratifying variable. The proportionality test on the model stratified by gender and D1 does not lead to a rejection of the hypothesis of proportionality (p-value=0.063). The estimates of the vector of parameter β are displayed in Table 4 (model-TC3 column).

Table 3 –Cox models testing the risk proportionality hypothesis.

	MODEL-C1		MODEL-C2		MODEL-C3	
	<i>chi-sq</i>	<i>p-value</i>	<i>chi-sq</i>	<i>p-value</i>	<i>chi-sq</i>	<i>p-value</i>
Gender (M)	13.70	<0.001	----	----	----	----
Sector (curing)	2.99	0.083	2.94	0.09	2.67	0.10
Sector(ISF)	0.98	0.321	0.77	0.38	0.45	0.50
Position (WC)	0.59	0.443	0.52	0.47	0.42	0.51
Position (BC)	2.15	0.142	2.24	0.13	2.15	0.14
BN(other)	0.47	0.491	0.40	0.62	0.17	0.68
BN(shipping company)	0.01	0.945	0.01	0.94	0.03	0.86
BN(COOP)	0.20	0.651	0.16	0.69	0.30	0.58
BN(corporate enterprise)	0.12	0.733	0.10	0.74	0.27	0.60
BN(private partnership)	0.08	0.778	0.09	0.76	0.32	0.57
Old	0.00	0.957	0.01	0.91	0.01	0.91
Age-start	0.56	0.452	0.76	0.38	0.96	0.33
D1 (True)	2.70	0.100	2.78	0.09	----	----
GLOBAL	45.24	<0.001	30.04	0.003	18.87	0.063

Note: BN: Business Name (“general company” has been used as reference category)

The only covariates representing the previous history are “age-start” and “old”. The first has a conservative effect: older people are more likely to obtain contracts with longer duration. Contrary to what could be expected, blue collars are more at risk than apprentices.

We have tried to improve the model introducing covariates describing the previous history, namely: “duration of previous unemployment”, “number of previous contracts”, “duration of previous unemployment>3 months”(Dis_3) and “duration of previous unemployment>12 months”(Dis_12) have been added to those of model TC3. Each time new covariates have been introduced the hypothesis of proportional hazard has been tested. The estimates for the new parameters have been depicted in Table 4 (model-TC1). None of the new covariates seems to have a significant effect on the duration of contract. Comparing the new model (TC1) with the previous (TC3) the Akaike Information Criterion (AIC) propends for the first model. Either the new covariates fail to summarize the previous history or the previous history does not influence the duration of future contracts.

Using the indexes introduced in the previous section yields a model (model-CY in Table 4) with a better AIC than model TC3. In the new model the covariates representing the previous history are the average of $da_{i,j}$ (da_m) and the standard

deviation of da_{ij} and dac_{ij} (da_sd , dac_sd). Model-CY shows a better AIC than the two previous models, and yet the covariates accounting for the previous history present p-values greater than 0.05. This could be due to the regressive nature of Cox model.

Table 4 –Cox model estimates

	MODEL-TC3		MODEL-TC1		MODEL-CY		MODEL-CZ	
	<i>Coeff.</i>	<i>p val</i>	<i>Coeff</i>	<i>p val</i>	<i>Coeff</i>	<i>p val</i>	<i>Coeff</i>	<i>p val</i>
Sector(curing)	0.17	0.09	0.17	0.10	0.17	0.09	0.17	0.08
Sector(ISF)	0.69	<0.01	0.70	<0.01	0.78	<0.01	0.77	<0.01
Position (WC)	0.02	0.91	0.01	0.96	0.14	0.32	0.13	0.38
Position (BC)	0.42	<0.01	0.42	0.02	0.44	0.01	0.43	0.02
BN(other)	0.17	0.68	0.18	0.67	0.28	0.50	0.26	0.53
BN(shipping company)	0.73	0.03	0.74	0.03	0.79	0.02	0.77	0.02
BN(COOP)	0.69	0.05	0.70	0.05	0.76	0.03	0.73	0.04
BN(corporate enterpr.)	0.73	0.03	0.74	0.03	0.84	0.01	0.81	0.02
BN(private partnership)	0.62	0.06	0.63	0.06	0.72	0.03	0.70	0.03
Old	-0.31	0.09	-0.36	0.08	-0.59	0.04	---	---
Age-start	-0.01	0.03	-0.01	<0.01	-0.01	<0.01	-0.01	0.01
Prev.Unempl.Time	---	---	-0.00	0.36	---	---	---	---
Num.of Prev.Contracts	---	---	-0.90	0.60	---	---	---	---
Dis_3	---	---	-0.02	0.88	---	---	---	---
Dis_12	---	---	-0.05	0.74	---	---	---	---
Dac_SD	---	---	---	---	0.29	0.85	---	---
Da_sd	---	---	---	---	-0.03	0.97	---	---
Da_m	---	---	---	---	-0.57	0.17	---	---
Cluster(1)	---	---	---	---	---	---	0.35	0.09
Cluster(2)	---	---	---	---	---	---	0.58	0.02
Cluster(3)	---	---	---	---	---	---	0.11	0.55
GLOBAL	18.87	0.06	24.21	0.06	20.89	0.10	18.83	0.13
AIC	6117.35	---	6124.13	---	6110.84	---	6108.53	---

Note: BN: Business Name ("general company" has been used as reference category)

Table 5 Cluster: previous history

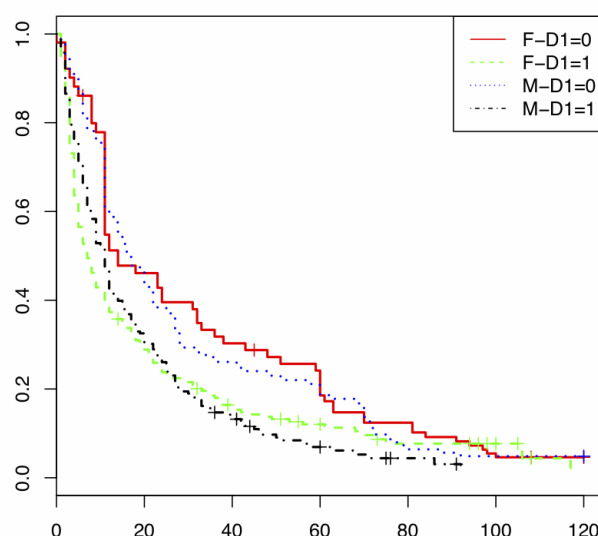
	Cluster1	Cluster2	Cluster3
Da_m	0.963	0.396	0.013
Da_SD	0.273	0.172	0.013
Dac_SD	0.034	0.203	0.015

We face this problem by looking for a clustering structure in the contracts with the history summarized as described above. Contracts already existing at 1st January 2000 have not been considered since for those contracts it is not possible to compute such quantities. A hierarchical clustering procedure based on the Ward

method has been used. A partition of the contracts into 3 clusters has been chosen. The totality of contracts is then partitioned into 4 clusters: the fourth being composed of those contracts already exiting at 1st January 2000.

Contracts belonging to cluster 1 are mainly of males, white collars, with an average number of previous contracts equal to 2.34 and an average unemployment period equal to 2.58 months. They show a high average value for $da_{i,j}$ (see Table 5) and a very small standard deviation, implying that they represent workers who are “continuously employed”. Contracts belonging to cluster 2 are mainly of males, with an average number of previous contracts equal to 3.11 and an average unemployment period equal to 7.4 months.

Figure 1 – Cox’s Model: stratified baseline hazard functions.



Contracts belonging to cluster 3 are mainly females apprentices with an average number of previous contract equal to 0.62 and average unemployment period equal to 38 months. They show (see Table 5) a very low values for the average and the standard deviation of $da_{i,j}$, implying that they represent workers who are “continuously unemployed”. A final model with the structural covariates and the cluster label has been fitted (Table 4, column model- CZ). This model shows the best AIC value. The cluster covariate is now significant in the model: namely those

contracts belonging to cluster 2 are more at risk than the contract in cluster 4 (used as corner point).

In Figure 1 the baseline hazard functions for each stratum have been displayed. According to non-proportional effects, baseline associated with different strata do intersect. Figure 1 clearly shows the relevance of variable D1: subjects with D1=1 have a higher risk of terminating the contract. Females present the best and the worst condition showing an interaction between the two factors.

5. Conclusions

The aim of this paper was to propose indicators to synthesize the pattern on the work history of an employee and to determine if such history influences the future duration of contracts. The results of the analysis show that variables accounting for previous work history influence the duration of contracts. This is highlighted in the behavior of baseline hazard functions associated with the values of variable D1 (see Figure 1) and in the significance of the cluster labels (a synthesis of the suggested indicators) that express the relevance of previous history (see Table 4).

Acknowledgments

The authors would like to thank the *Osservatorio Nazionale della Pesca* for providing the data used in the analysis.

Reference

THERNEAU T.M. GRAMBSCH P.M., 2000, Modeling Survival Data: extending the Cox model (Statistics for biology and health), Springer

SUMMARY

Determinants of the duration of contracts: a longitudinal study in the Italian fishery sector

The aim of this work is to summarize the previous work history of an employee and to evaluate if it has an influence on the duration of employment. We have analyzed the reality of Italian contracts in the years from 2000 to 2009 focusing the study on a random sample of contracts from the fishery sector.

Luciano NIEDDU, Libera Università LUSPIO, l.nieddu@luspio.it

Cecilia VITIELLO, "Sapienza" Università di Roma, cecilia.vitiello@uniroma1.it