TÍTULO

JOB RETENTION SCHEMES AND IMPLICANTIONS ON FIRM SURVIVAL DURING COVID-19. EVIDENCE FROM NEW SPANISH DATA

AUTOR

Javier García Clemente

Esta edición electrónica ha sido realizada en 2022

Tutores	D. Emilio Congregado ; D. Jesús Rodríguez
Instituciones	Universidad Internacional de Andalucía ; Universidad de Huelva
Curso	*Máster en Economía, Finanzas y Computación (2020/21)*
©	Javier García Clemente
©	De esta edición: Universidad Internacional de Andalucía
Fecha documento	2021

Job retention schemes and implications on firm survival during Covid-19. Evidence from new Spanish data

by

Javier García Clemente

A thesis submitted in conformity with the requirements for the
MSc in Economics, Finance and Computer Science

University of Huelva & International University of Andalusia

November 22, 2021

Job retention schemes and implications on firm survival during Covid-19. Evidence from new Spanish data

Javier García Clemente

Supervisor: Emilio Congregado & Jesús Rodríguez

Abstract

We analyze the aggregated survival rates of more than a million employer units followed quarterly from April 1st, 2020 to April 1st, 2021, using new data from the experimental statistics of Demographic Situation of Companies (CODEM), by the Spanish National Statistics Institute (INE). This approach makes use of fractional regression methods to explain the survival rate by region, sector, size of the company and whether or not a job retention scheme (henceforth ERTE) had been adopted. Our main results show that the survival rate were significantly higher in those companies where an ERTE was adopted. Nevertheless, this effect was not homogeneous, particularly benefiting the most vulnerable firms. These firms were, as expected, the smallest (from 1 to 5 employees) and the ones which operate in some service sectors related to leisure, education, tourism and hospitality. Additionally, some unobserved heterogeneity among regions have been considered too, suggesting Balearic and Andalusian firms to be the most likely to close.

JEL classification: E65, L20, L50, D20, H32.

Keywords: Firm survival, Covid19, Job Retention Scheme, ERTE, business closings.

Resumen

En este trabajo analizamos las tasas de supervivencia agregadas de más de un millón de unidades empleadoras, seguidas trimestralmente desde abril de 2020 hasta abril de 2021, con datos españoles de la nueva estadística sobre Situación Demográfica de Empresas (CODEM), del Instituto Nacional de Estadística (INE). Para ello, utilizamos métodos de regresión fraccional con el objetivo de explicar la tasa de supervivencia empresarial por regiones, sectores, tamaño de empresa y si estas se acogieron a un Expediente de Regulación Temporal de Empleo (ERTE). Los principales resultados muestran que la tasa de supervivencia fue significativamente mayor en aquellas empresas que adoptaron el ERTE. Sin embargo, dicho efecto parece no ser homogéneo, beneficiando especialmente a las empresas más vulnerables. Por su parte, observamos que estas empresas más vulnerables fueron las más pequeñas (de 1 a 5 empleados), y aquellas que operaban en servicios relacionados con el ocio, la educación, el turismo y la hostelería. Por último, la heterogeneidad a nivel regional sugiere una menor tasa de supervivencia empresarial en determinadas Comunidades Autónomas, destacando Baleares y Andalucía.

Key findings

- Education, leisure-related services and micro-enterprises experienced the lowest survival rates during the pandemic.
- On the other hand, industries, health care services and larger businesses were the most resilient profiles.
- Firms whose employees were sent to ERTE schemes are significantly associated with higher survival rates, controlling by their main characteristics.
- However, this ERTE positive effect seems to be greater for the weakest firms and smaller for the resilient, working as a sort of convergence effect.
- The regional unobserved heterogeneity point at Northern Spain located firms as the most surviving ones, in contrast to Balearic and Andalusian companies, the least surviving.

Contents

List of Figures

List of Tables

1. Introduction

The survival of the companies is one of the multiple socioeconomic dimensions which have been negatively shocked by the pandemic. From an economic perspective, this is one of the most important issues which governments have dealt with, since the generalized lockdown and the social distancing measures led to the sudden closing of many firms. Nonetheless, despite being generalized throughout the whole economy, the impact of this shock may not have been equal for all companies. In this regard, the economic sector or the company size may have been crucial when talking about the survival of the companies throughout the pandemic.

In addition, with the aim to relieve the negative impact of the lockdown in the labor market, some kind of job retention scheme was massively used in most countries during the early stages of the sanitary crisis, as reported in OECD (2020) and Drahokoupil and Müller (2021). The latter even makes a distinction for job retention schemes in three categories: short-time work schemes, furlough schemes and wage subsidies. Nevertheless, as mentioned by the authors, there is some overlap among those terms, so we will use them indistinctly throughout this paper.

Focusing on Spain, this policy (known as ERTE) consists in a temporary suspension of the labor relationship between the employer and the employee, or alternatively, a reduction of working hours justified by a major cause. This cause must be related to economic, technical, organizational or production issues, including Covid related consequences from March, 2020. During this period of suspension, the employee is getting a social security allowance while the employer only has to assume a social contribution, which is a minor part of the employee´s wage. As a result, it works as a transitory mechanism of flexibility to adjust the labor market, whose cost is essentially assumed by the public administration. This policy main purpose is to maintain the employees position despite not being working, avoiding a sharp boost of unemployment during the shock. However, since the firms are not assuming the adjustment costs (labor or redundancy costs), this policy may avoid business closings and help them overcome transitory shocks via the introduction of labor demand flexibility.

Despite existing before the pandemic, the Spanish ERTEs were only widely used then, covering around 3 million of workers (more than 20% of the affiliated workers) in the second quarter of 2020. The following quarters it covered around 5% of the affiliated workers, which is still a significantly higher proportion than it was during the previous recession. (Izquierdo Peinado et al., 2021)

Moreover, since the Spanish government first approved these Covid-19 related ERTEs in March, 2020, their expiration date has been postponed several times, remaining in the current legislation. Therefore, some evaluation of the impact of this policy in all dimensions is urgently needed to improve the design of these programs for the next future. Up to now, some authors have tried to do this evaluation but focusing on employment and labor market implications. By contrast, we have focused our question this time on whether or not this job retention scheme may have braked the business exit, and if so, in what extent.

Furthermore, the delay in the traditional data release have forced many researches to seek alternative data sources in the midst of this unprecedented shock in order to measure the impact of the pandemic. Nonetheless, to the best of our knowledge no studies have been published using this Spanish Demographic Situation of Companies database so far.

To sum up, our main contribution comes from the use of this new database to measure the heterogeneity in the firm survival rate in Spain during the pandemic. Additionally, as these data is linked to the information about job retention scheme adoption, we also contribute to the literature by finding an association of the Spanish ERTE policy with firm survival rates.

The next pages are organized as follows: In section 2 we review some of the related literature; section 3 introduces the database and methods; section 4 contains a descriptive analysis followed by the main results of the regressions performed; section 5 summarize the main conclusions about the results; and finally, a discussion about the contributions and limitations is presented in section 6.

2. Related literature

At the beginning of the pandemic, Bartik et al. (2020) conducted a survey to quantify the early impact of the Covid-19 outbreak on US small businesses, finding massive layoffs and shutdowns as a result. Similarly, Fairlie (2021) observed an unprecedented drop in active business owners in the early 2020 using US Current Population Survey data. Shortly afterwards, Fairlie and Fossen (2021) analyzed the huge drops in business sales during the second quarter of 2020, using administrative data from California. However, these effects seem to be very heterogeneous: According to the latter study, sale losses were largest in business affected by mandatory lockdowns such as accommodation (-91%), in contrast to other businesses which experienced large gains, such as online sales (+180%). By the same token, Crane et al. (2020) discovered that the US business exit during the early outbreak was elevated for certain sectors, but lower than usual for other industries. Bloom et al. (2021), despite reporting overall negative impacts of Covid-19 on US firms sales, observed that it depends significantly on the firm size, the percent of the online revenue and the business owner´s demographics. Additionally, using small business data from Oakland, Bartlett and Morse (2020) found noticeable differences in the firm survival capability during the pandemic, especially by its size. These differences, also affected the effectiveness of business assistance programs, as the US Payroll Protection Program (PPP). So far, the main contribution to the topic using Spanish data come from Fernández Cerezo et al. (2021). They use business data from a Bank of Spain own survey to measure variations in company sales and employment, finding great heterogeneity mainly among sectors and firm sizes.

By contrast, most of the Short Time Work Scheme (henceforth STWS) literature use pre-Covid data and do not focus its attention on the implications for firms, but for employees and its capability to retain them. A seminal example come from Hijzen and Venn (2011), who designed a natural experiment for 19 OECD countries in order to estimate the causal effects of STWS on preserving jobs in the Great Recession, finding overall but not widespread positive effects. Afterwards, in Hijzen and Martin (2013), they remarked the crucial role of good timing in the application of these public policies.

More recently, Cahuc (2019) explored the pros and cons of STWS, highlighting them as a worthy instrument to temporarily adjust the labor market during recessions, but also alerts about the risk of turning to an inefficiency source. However, the most recent and remarkable contribution can be found in Cahuc et al. (2021). Using French data they measured the heterogeneous effects of STWS at saving jobs and increasing total hours of work considering the Great Recession period. Their results pointed out the effectiveness of STWS for firms hit by strong negative revenue shocks. On the contrary, for firms facing a limited decrease in revenues this policy only would have been a sort of "windfall effect", in authors´ words.

3. Data and Methodology

3.1. Data

In order to analyze the survival of the companies during the first year of the Covid-19 pandemic in Spain we retrieved data from the new experimental statistics of Demographic Situation of Companies (CODEM), by the Spanish National Statistics Institute (INE).[1] These data gather aggregated information about the number of surviving companies and its rate regarding a initial cohort of 1,102,738 employer units followed from April 1st, 2020 to April 1st, 2021, in a quarterly frequency.

To understand the nature of the database, it consists in an aggregated panel where we are monitoring the fraction of surviving companies from the initial cohort. The cross-sectional dimension is defined by a given region, ERTE situation and sector or size category, while observed in each of the 4 following periods.

The main limitations of these data comes from two sides. On one hand, we cannot access the firm-level microdata with the 1,102,738 observations, accessing only to its aggregation by sector or size instead. On the other, we cannot combine the sector dimension with the firm size because the data is provided separately. As a consequence, we will be simultaneously conducting an analogous analysis for both datasets, the first, considering the sector aggregation, and the second, considering the size aggregation. In spite of these limitations, the main point of this new database is the possibility to link the information about the adoption of job retention schemes with firm survival rates, allowing the introductory analysis of an hypothetical relationship between them.

[1] `https://ine.es/experimental/codem/experimental_codem.htm` (last consult July 5th, 2020).

Hence, the first dataset (henceforth referred as Sector aggregation) contain aggregated data for 4 quarters, 17 regions (Spanish Autonomous Communities), 81 sectors of the National Classification of Economic Activities (CNAE) which were aggregated into 16 sectors for interpretive purposes and 2 ERTE situations (yes or no). Considering 12 missing values it comes to 2,164 observations altogether. Then, we were forced to drop the sectorial dimension to include a categorical number of employees, as a proxy of the company size, for the second dataset (henceforth referred as Size aggregation), leading to 680 observations with no missing values. Thus, any observation represents the rate of companies which did not close over the initial cohort, observed in a given period, region, ERTE situation and either sector or size, as shown in the example of tables 1 and 2.

Table 1: Example for data visualization in the sector aggregated dataset.

Region	Sector	Erte	t	Survival rate
Andalucía	Extractive industry	0	1	.9406
Andalucía	Extractive industry	0	2	.9300
Andalucía	Extractive industry	0	3	.8812
Andalucía	Extractive industry	0	4	.8554
...	...	...	...	...
La Rioja	Other services	1	1	.9861
La Rioja	Other services	1	2	.9861
La Rioja	Other services	1	3	.9535
La Rioja	Other services	1	4	.9210

Data source: CODEM.

Table 2: Example for data visualization in the size aggregated dataset.

Region	Size	Erte	t	Survival rate
Andalucía	From 1 to 5	0	1	.9373
Andalucía	From 1 to 5	0	2	.8894
Andalucía	From 1 to 5	0	3	.8427
Andalucía	From 1 to 5	0	4	.8005
...	...	...	...	...
La Rioja	More than 250	1	1	1
La Rioja	More than 250	1	2	.8333
La Rioja	More than 250	1	3	.8333
La Rioja	More than 250	1	4	.8333

Data source: CODEM.

3.2. Methodology

Following by a first descriptive observation of the data which suggested differences in the survival rate by sector, size, region and ERTE situation, some fractional logit regressions with robust default standard errors were carried out in STATA 16[2] to estimate the relationship of every of these characteristics with the survival rate. The choice of this technique is justified by the fractional nature of the outcome variable, survival rate. As it takes continuous values within the close interval [0, 1], the effect of any particular covariate cannot be constant throughout the range of the vector x and the predicted values from an OLS regression can never be guaranteed to lie in the unit interval. In addition, by using this technique we can relax some of the assumptions of the linear regressions regarding linearity, normality and homocedasticity.

This fractional regression method, initially proposed by Papke and Wooldridge (1996), simply solves these issues using a quasi-likelihood estimator and a logistic distribution for the conditional mean of the dependent variable. Therefore, the log-likelihood function for these fractional models is of the form of equation 1.

$$lnL = \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it} ln \left\{ G(x_{it}^{'}\beta) \right\} + (1 - y_{it}) ln \left\{ 1 - G(x_{it}^{'}\beta) \right\} \tag{1}$$

where lnL is maximized, N denotes the cross-section sample size, T is the time dimension, y_{it} is the dependent variable (for us, the survival rate), $G(.)$ represents a logistic functional form (thus $exp(x_{it}^{'}\beta)/\{1 + exp(x_{it}^{'}\beta)\}$) and x_{it} are the covariates for any observation i in period t. Finally, the β denotes the parameters to be optimized by numerical algorithms.

As mentioned, the conditional mean for the survival rate will follow the logistic distribution given by equation 2.

$$E(y_{it}|x_{it}) = exp(x_{it}^{'}\beta)/\{1 + exp(x_{it}^{'}\beta)\} \tag{2}$$

3.2.1. 1st regression: sector aggregated data

The set of covariates x included in the first regression of the company survival rate were all dummy variables by sector, regions, quarters and ERTE. These covariates may allow us to catch the heterogeneity in the company survival rate among sectors, regions, time and finally the ERTE estimated effect via the computed marginal effects of each covariate. Note that since all the covariates are factor levels, the marginal effects will represent discrete changes from the base level. In this specification, the omitted categories were hospitality and tourism sector, trend dummy=1, Andalusia and no ERTE.

[2] Original codes are available in the appendix to ensure the reproducibility of the results.

3.2.2. 2nd regression: size aggregated data

Likewise, the second regression was based on the same structure as the previous one, but replacing the sector dummies by the size dummies. Then, the marginal effects may differ from the first regression and the size effects on survival rate will be quantified too. For this regression, the omitted categories were 1-5 employees size, trend dummy=1, Andalusia and no ERTE.

4. Results

4.1. Descriptive statistics

We are considering the survival rate as the dependent variable. For both datasets its values are bounded between 0 and 1 but its frequency distribution is skewed to the upper values. It is explained because at the initial period (t=0), all observations begin with survival rate = 1 (all companies alive), then it starts decreasing gradually (when some companies close) for the subsequent periods (t=1,2,3,4). Table 3 displays the main descriptive stats of the survival rate for both aggregated datasets, considering all observations except for the initial period (t=0), which have been dropped because it would not provide any useful information for the aforementioned reason. There, we observe survival rate means above 90% in both cases and median values even greater. In both cases, the majority of the distribution is over 90%, as shown by their percentiles. Just to clarify, since the two datasets are aggregating the companies in a different way, the distribution of the survival rates may not be equal for both, as it happens.

Table 3: Survival rate (S) unconditional descriptive stats.

	mean	sd	min	max	p25	p50	p75	N
S, sector aggr.	0.928	0.065	0.606	1	0.898	0.943	0.976	2164
S, size aggr.	0.967	0.043	0.738	1	0.957	0.982	0.995	680

Data source: CODEM.

In order to identify the potential relationship between the survival rate and the sector, size, region and ERTE situation, we analyzed the conditional distribution of the survival rate by these attributes. To sum up these results, some boxplots will be displayed in the following pages. These kind of graphs are useful to provide a great amount of information about the distribution of a numeric variable at a glance, especially when establishing a comparison over different groups. As a brief explanation for someone unfamiliar with them, the boxs represent the interquartile range, thus, the range between the first and fourth quartile of the variable distribution (survival rate for us). The line subdividing that box represents the second quartile or equivalently, the median of the distribution. The lines laying outside the box, often called whiskers, extends to the furthest data point in each wing that is within 1.5 times the interquartile range.

Finally, any data point further than that distance is considered an outlier, and is marked with a dot. Additionally, note that the over group comparison will be displayed in a descending order by median, just for an easier interpretation.

Figure 1, which shows the survival rate regarding the ERTE situation, reveal higher survival rates for firms in ERTE. Using the sector aggregated dataset, the survival rate mean is 90.21% for non-ERTE companies and 95.33% for ERTE´s ones. Nevertheless, this difference seems to be lower in the size aggregated dataset (96.21% survival mean for non-ERTE companies and 97.11% for ERTE´s).

Moreover, in figure 2 some regional heterogeneity can be observed in both datasets, although it does not seem high. The survival rate mean fluctuate around 2 percentage points above and below regarding the region. On the other hand, according to figure 3, the survival rate of the cohort seems to fall proportionally every quarter.

In addition, figures 4 and 5 reveal interesting patterns for the survival rate. Particularly, in figure 4, the existing heterogeneity among sectors is clear, remarking education (82.33% survival rate mean), arts/entertainment (88.37%) and hospitality/tourism (89.96%) as the most vulnerable sectors, as oppose to extractive (95.91%) and supply industries (96.46%), which were the most resilient.

Lastly, the survival rate seems to be directly related with the company size, increasing the survival as the company is larger. (see Figure 5). In firms with less than 5 employees the survival mean was only 90.58%, while firms with 6-9 employees increased this mean up to 96.68%, and larger firms up to 98 or 99%.

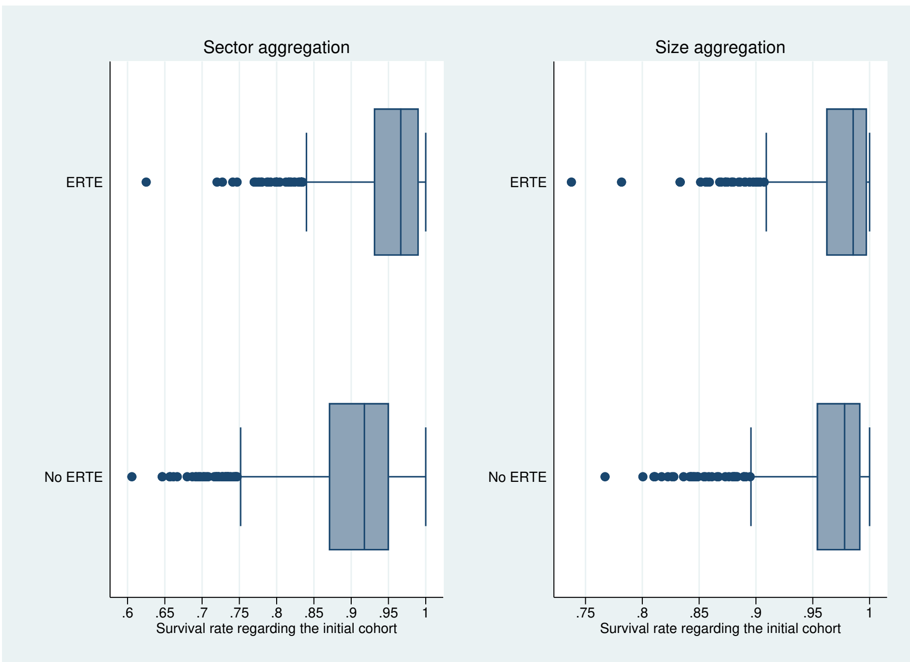

Figure 1: Survival rate by ERTE situation. Data source: CODEM.

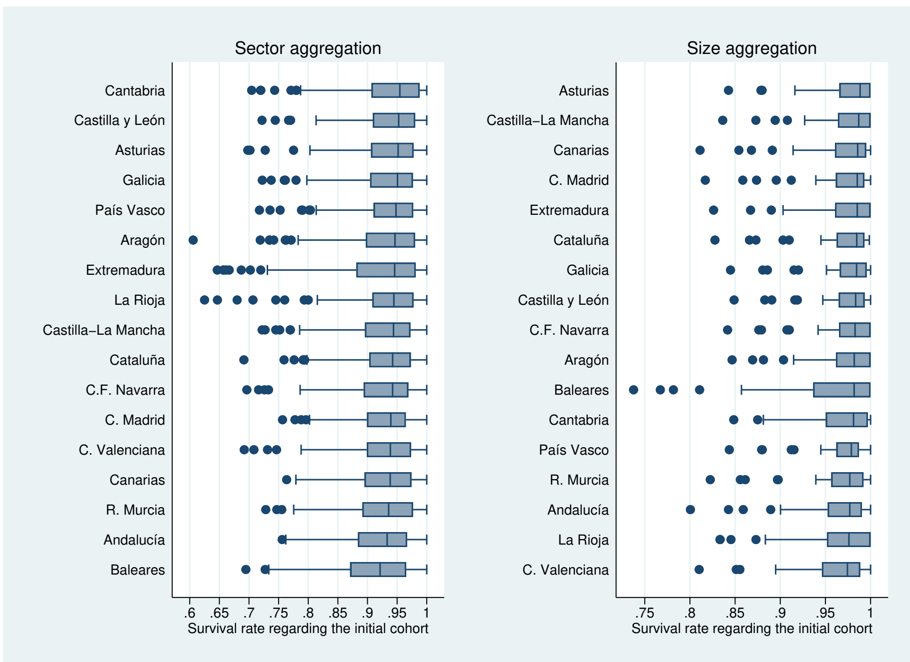

Figure 2: Survival rate by region. Data source: CODEM.

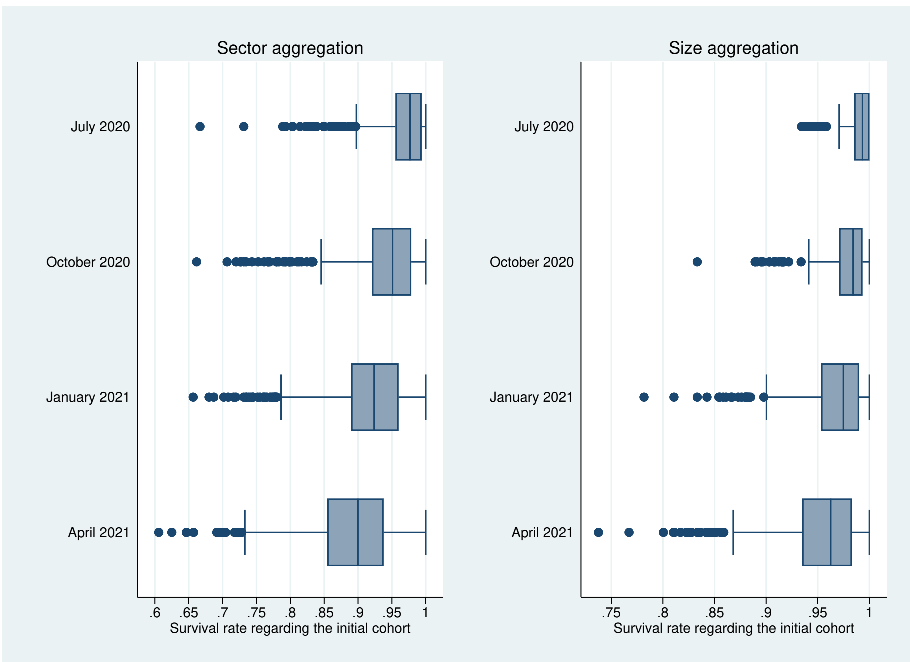

Figure 3: Survival rate by quarters since the initial period (April 1st, 2020). Data source: CODEM.

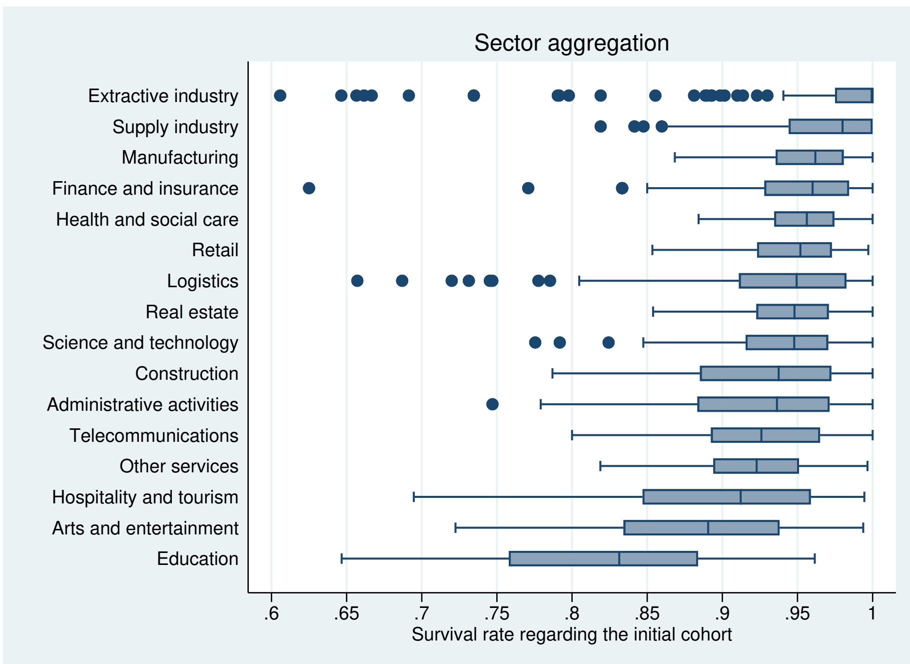

Figure 4: Survival rate by sector. Data source: CODEM.

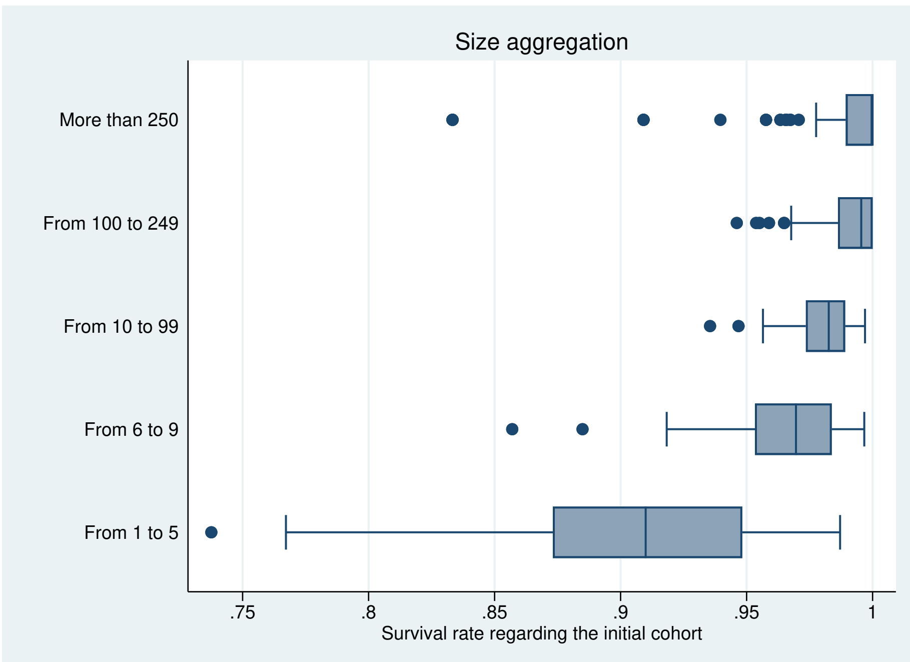

Figure 5: Survival rate by size (number of employees). Data source: CODEM.

4.2. Regressions

As the fractional logit regression output is not directly interpretable, we are focusing on their estimated marginal effects. Table 4 shows the average marginal effects for both regressions. All these effects have been tested significant except for some regional ones and represent the discrete change from the baseline survival outcome since they are associated with factor variables. Thus, multiplied by 100, they can be interpreted as increments or decrements percentage points in baseline survival rate.

Hence, on average, taking up an ERTE increased the survival rate by 5.31 percentage points (henceforth shorten as pp) in the first regression. Nonetheless, in the second regression this effect is atenuated to less than 1pp. Furthermore, the trend effects seems to follow a linear decreasing pattern, proportionally reducing the survival every quarter. The sector average marginal effects confirm the hypothesis of the most vulnerable sectors (education, hospitality/tourism and arts/entertainment), highlighting education as the less likely to survive (7.6pp below hospitality and tourism). On the other hand, extractive and supply industries and health/social services were the most resilient (more than 5pp over hospitality/tourism). In addition, the size effect estimated in the second regression tell us that a company with 6 to 9 employees increases the survival rate by more than 6pp with regard to a 1 to 5 employees firm. However, this increment is marginally declining as the company is larger. Finally, the regional effects point out the expected heterogeneity among regions, highlighting Andalusia and the Balearic Islands as the least surviving regions for the firms.

4.2.1. Deepening into ERTE marginal effect

Although we have found a positive ERTE effect in both regressions, we have only computed the marginal effect on average, this is, considering the average marginal effect over all possible values of the rest of covariates. Being a nonlinear regression, the marginal effect of a given covariate may differ across the values of other covariates. Fixing values of other covariates we can estimate the variation of this effect at different points, interacting the ERTE effect with them. From now on, we are going to explore the marginal effect of the ERTE at different values for time, sector and size. This analysis would reveal if the average effects which have been obtained in table 4 are stable over time, sectors and size. As showed by figure 6, the ERTE effect is not constant over time, increasing proportionally every quarter. This result suggests that the gap in survival rates between ERTE and non ERTE companies widened every quarter. By sector, the ERTE positive effect seems to be larger for the least surviving sectors, thus education, hospitality/tourism and arts/entertainment. The huge difference in the ERTE positive effect on survival rates by sector reach almost 10pp, from +2.8pp in the supply industry to +11.4pp in education (Figure 7). By size, these results are quite relevant as well since the ERTE positive effect lay between 2-3pp of survival rate increase only for the smallest firms, decreasing sharply near zero when the size of the company is greater (Figure 8).

Table 4: Average marginal effects on survival rate.

	(reg. 1) Sector aggregation		(reg. 2) Size aggregation	
ERTE	0.0531	(0.000)	0.0090	(0.000)
Trend dummy=2	-0.0263	(0.000)	-0.0143	(0.000)
Trend dummy=3	-0.0521	(0.000)	-0.0289	(0.000)
Trend dummy=4	-0.0778	(0.000)	-0.0435	(0.000)
From 6 to 9 employees			0.0609	(0.000)
From 10 to 99 employees			0.0747	(0.000)
From 100 to 249 employees			0.0856	(0.000)
More than 250 employees			0.0824	(0.000)
Extractive industry	0.0611	(0.000)		
Manufacturing	0.0565	(0.000)		
Supply industry	0.0652	(0.000)		
Construction	0.0259	(0.000)		
Retail	0.0449	(0.000)		
Logistics	0.0316	(0.000)		
Telecommunications	0.0228	(0.000)		
Finance and insurance	0.0493	(0.000)		
Real estate	0.0452	(0.000)		
Science and technology	0.0393	(0.000)		
Administrative activities	0.0266	(0.000)		
Education	-0.0764	(0.000)		
Health and social care	0.0546	(0.000)		
Arts and entertainment	-0.0160	(0.000)		
Other services	0.0214	(0.000)		
Aragón	0.0078	(0.095)	0.0079	(0.000)
Asturias	0.0170	(0.000)	0.0123	(0.000)
Baleares	-0.0105	(0.008)	-0.0107	(0.015)
Canarias	0.0084	(0.006)	0.0062	(0.005)
Cantabria	0.0131	(0.007)	0.0025	(0.537)
Castilla y León	0.0174	(0.000)	0.0093	(0.000)
Castilla La-Mancha	0.0077	(0.028)	0.0113	(0.000)
Cataluña	0.0087	(0.020)	0.0063	(0.001)
C. Valenciana	0.0053	(0.058)	-0.0034	(0.189)
Extremadura	-0.0053	(0.397)	0.0088	(0.000)
Galicia	0.0143	(0.000)	0.0101	(0.000)
C. Madrid	0.0090	(0.002)	0.0051	(0.009)
R. Murcia	0.0038	(0.373)	0.0016	(0.490)
C. F. Navarra	0.0051	(0.197)	0.0080	(0.000)
País Vasco	0.0120	(0.000)	0.0043	(0.036)
La Rioja	0.0086	(0.063)	-0.0046	(0.533)
Survival baseline outcome	0.9289		0.9557	

p-values in parentheses.

Baseline categories: ERTE=0; Trend dummy=1; 1-5 employees;
Hospitality and tourism; Andalucía.

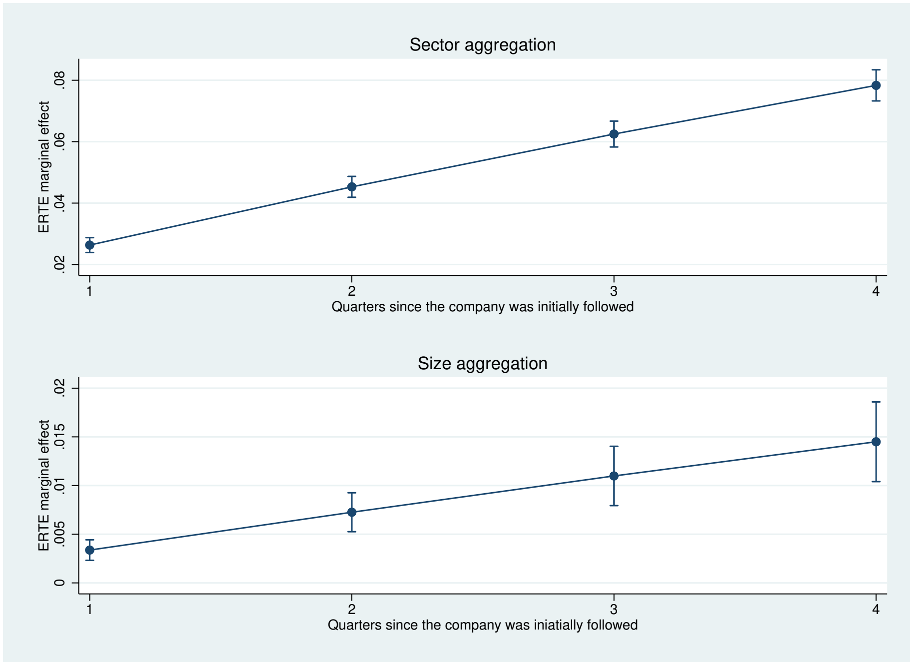

Figure 6: Erte marginal effects on the conditional mean of the survival rate with 95% CIs, by quarter.

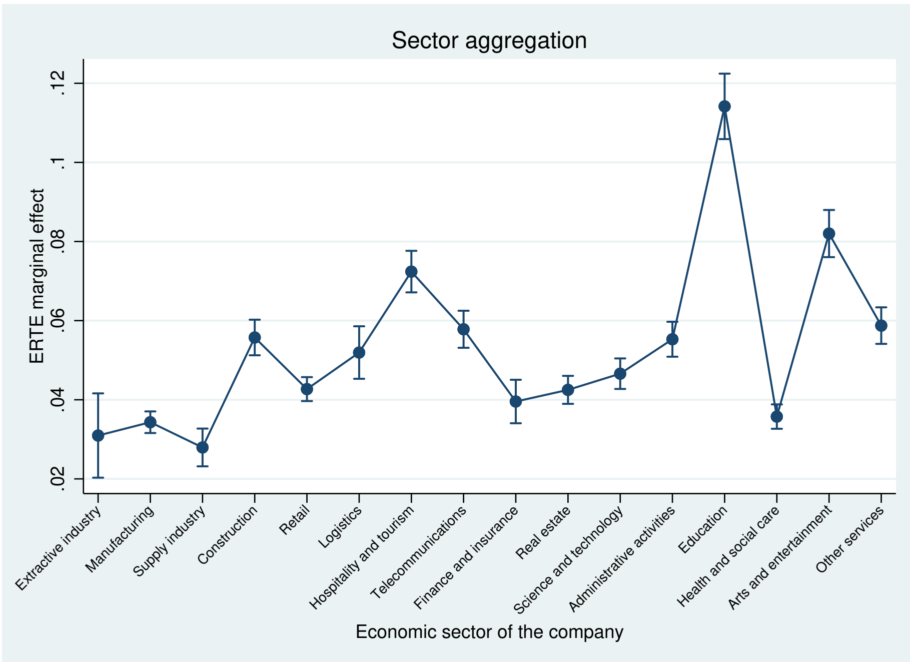

Figure 7: Erte marginal effects on the conditional mean of the survival rate with 95% CIs, by sector.

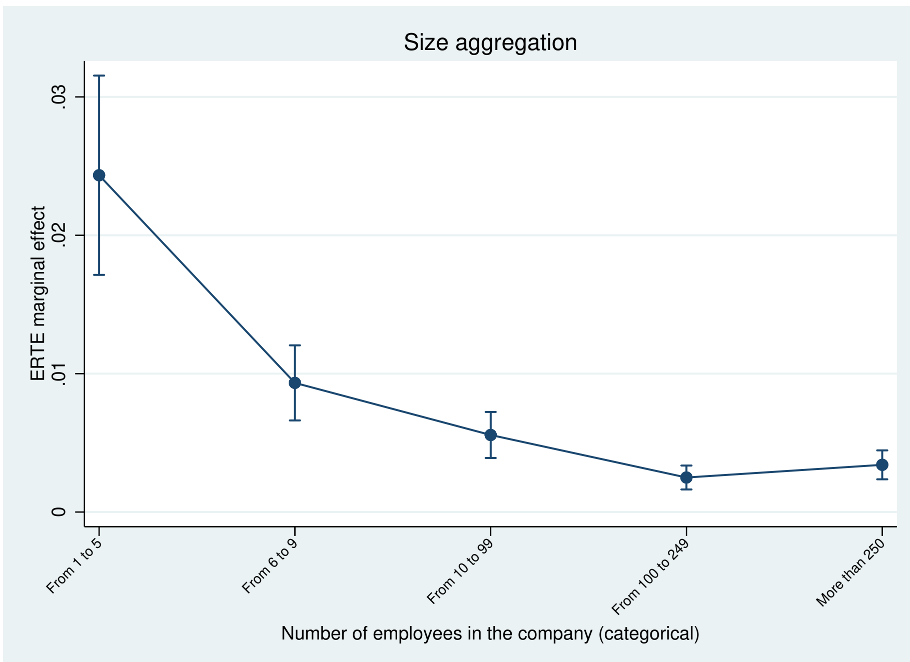

Figure 8: Erte marginal effects on the conditional mean of the survival rate with 95% CIs, by size.

5. Conclusions

Along this paper we have analyzed the survival rate of companies during the first year of Covid-19 in Spain, trying to measure the heterogeneity among sectors, size, regions and quantifying somehow the ERTE effects on it. For this purpose, we made use of the Spanish Demographic Situation of Companies new database, which imply a methodological contribution to this topic. Then, the main analysis consisted in some fractional logit regressions, performed in two different datasets due to the limitations for combining variables in the aggregated data provided by the source.

The first regression tried to estimate the survival heterogeneity among sectors; remarking education, arts and entertainment, hospitality and tourism as the sectors with less survival, while industries and health care had the highest survival rates, with a gap of 14pp on average from the lowest to the highest. Moreover, a positive survival effect for companies in ERTE situation was detected. In addition, we observed that this effect increased across time and it was significantly different among sectors, laying from +2.8pp in the supply industry to +11.4pp in education. This result, suggest a sort of convergence positive effect of ERTE since it has benefited more the vulnerable sectors, despite also benefiting the strongest sectors to a lesser extent. Regarding the regional effect, the unobserved heterogeneity caught by the regional dummies suggested higher survival rates in Northern regions and less survival in Baleares and Andalucía.

Likewise, the second regression quantified the heterogeneous survival by size, highlighting that in the smallest companies (less than 5 employees) the survival rate was on average 6pp lower than 6-9 employees companies. This gap raised up to more than 8pp comparing the smallest with the largest, with more than 250 employees. Nonetheless, the increments in survival rate by size seems to be marginally decreasing, finding a noticeable leap between 1-5 and 6-9 employees firms, but less relevant leaps for the subsequent sizes.

With regard to the ERTE positive effect in the second regression, it seems to substantially depend on the size too. Whereas in the smallest companies the ERTE increased the survival by 2 or 3pp on average, this effect dropped below 1pp in companies with more than 6 employees.

6. Discussion

Our results about firm survival in Spain during the pandemic point at the same direction as the reviewed literature for the US, this is, there was great heterogeneity and a significant part of it can be explained by the sector and size of the company. On the other hand, our characterization of the effect of a job retention scheme on firm survival by different type of firms suggests that its heterogeneity should be taken into account when designing these job retention mechanisms. Moreover, the association found between job retention schemes and firm survival opens the door for the study of a potential causal relationship.

Nevertheless, despite obtaining some interesting results, we should be cautious in their interpretations. Firstly, it should be reminded that an ERTE might be approved only for few employees within a company and also it could be either part or full time. Hence, the different characteristics of this policy may biased this result, since we are assuming an homogeneous ERTE scheme. Furthermore, we found some other troubles we could not overcome in the analysis: we suspect the ERTE adoption may not be exogenous at all since companies with less survival expectations might be more likely to adopt an ERTE at some point than the others; and also there might be relevant omitted variables in our models which cannot have been included due to the few control variables available. In addition, the techniques we have used are not able to infer any causality effect beyond correlation. Therefore, some public evaluation techniques, e.g. the use of counterfactual scenarios, are needed in order to properly evaluate these effects. On the other hand, using microdata, other techniques like pure survival analysis with duration models would perform better to estimate the survival of any company controlling by its characteristics. Unfortunately, none of these approaches have been possible this time with the available aggregated data.

Being aware of all these limitations, this paper is only aiming to introduce some preliminary insight on the Covid-19 impact in the company survival in Spain, as well as provide useful information about the possible role of the job retention schemes in this topic. As far as we are concerned, further research is needed as it is a new topic with few available data, but of utmost importance nowadays. We will keep working on the following issues for future research: the use of firm-level weights to improve the regressions performance and other estimation methods; the comparison of these firm survival rates during the pandemic with pre-pandemic data in order to confront these data with a baseline survival reference; the possibility of merging these data with other sources to consider new variables in the analysis which might explain the regional survival heterogeneity, e.g. industry composition shares or bureaucratic and transaction costs for companies; and finally, the design of a causal evaluation for job retention schemes.

References

Bartik, A. W., M. Bertrand, Z. B. Cullen, E. L. Glaeser, M. Luca, and C. T. Stanton (2020). How are small businesses adjusting to COVID-19? Early evidence from a survey. Technical report, National Bureau of Economic Research (NBER).

Bartlett, R. P. and A. Morse (2020). Small business survival capabilities and policy effectiveness: Evidence from Oakland. Technical report, National Bureau of Economic Research (NBER).

Bloom, N., R. S. Fletcher, and E. Yeh (2021). The impact of COVID-19 on US firms. Working Paper 28314, National Bureau of Economic Research (NBER).

Cahuc, P. (2019). Short-time work compensation schemes and employment. *IZA World of Labor* (11).

Cahuc, P., F. Kramarz, and S. Nevoux (2021). The heterogeneous impact of short-time work: From saved jobs to windfall effects. *IZA Discussion Paper* (14381).

Crane, L. D., R. A. Decker, A. Flaaen, A. Hamins-Puertolas, and C. Kurz (2020). Business exit during the COVID-19 pandemic: Non-traditional measures in historical context. Technical report, Finance and Economics Discussion Series (FEDS).

Drahokoupil, J. and T. Müller (2021). Job retention schemes in europe: a lifeline during the Covid-19 pandemic. *ETUI Research Paper-Working Paper*.

Fairlie, R. and F. M. Fossen (2021). The early impacts of the covid-19 pandemic on business sales. *Small Business Economics*, 1–12.

Fairlie, R. W. (2021). The impact of COVID-19 on small business owners: The first three months after social-distancing restrictions. Technical report, National Bureau of Economic Research (NBER).

Fernández Cerezo, A., B. González, M. Izquierdo Peinado, and E. Moral Benito (2021). El impacto económico del COVID-19 en las empresas españolas según la encuesta del banco de españa sobre la actividad empresarial (EBAE). *Boletín económico del Banco de España* (1).

Hijzen, A. and S. Martin (2013). The role of short-time work schemes during the global financial crisis and early recovery: a cross-country analysis. *IZA Journal of Labor Policy 2*(1), 1–31.

Hijzen, A. and D. Venn (2011). The role of short-time work schemes during the 2008-09 recession. *OECD Social, Employment and Migration Working Papers. OECD Publishing.* (115).

Izquierdo Peinado, M., S. Puente Díaz, and A. V. Regil Amorena (2021). Los erte en la crisis del covid-19: un primer análisis de la reincorporación al empleo de los trabajadores afectados. *Boletín económico del Banco de España* (2).

OECD (2020). *Job retention schemes during the COVID-19 lockdown and beyond*. OECD Publishing.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics 11*(6), 619–632.

A. Appendix

STATA CODES

```
* DO FILE - COMPANY SURVIVAL DURING COVID19
**********************************************************
*** IMPORT FROM EXCEL FILES AND PREPROCESSING ***
**********************************************************

****************************************************
** DATABASE 1: BY REGIONS, SECTORS AND ERTE **
****************************************************
clear all
set more off
capture log close
cd "C:WorkingDirectory" // Set your working directory

** APRIL 2020 COHORT DATA **
* Setting initial values of global macros for the loop
global inicio 6
global final 252
global i 0
while $i < 20 {

   import excel "C:WorkingdirectoryAndFileName.xlsx", sheet("1 ABRIL 2020")
   cellrange(A$inicio :K$final ) firstrow case(lower) // import code
   * format of survival rates *
   local S s1 s2 s3 s4
   foreach s in `S' {
     destring `s', float replace
     replace `s' = `s'/100 // survival in fraction instead of percentage
     }
   * generating regional datasets *
   reshape long q s, i(cnae erte) j(t) // reshape to long data
   gen cohort = 1 // id for cohort
   gen ccaa = 0 + $i // id for region
   replace s = 1 if t==0 // survival rate in the initial period == 1 always
   save "data_cohort1_region$i.dta", replace
   clear
   * iterating the loop *
   global i = $i + 1
   global inicio = $inicio + 250
   global final = $final + 250
   }
```

```
* Appending regional datasets
use "data_cohort1_region0.dta", clear // master dta file
forval n = 1/19 {
    append using "data_cohort1_region‘n’.dta" // append loop
    erase "data_cohort1_region‘n’.dta" // erasing unnecessary files
    }
compress
save "data_sectors.dta", replace // saving database
erase "data_cohort1_region0.dta"

* Aggregating sectors
gen sector = 0 if cnae==0
replace sector = 1 if cnae >=5  cnae < 10
replace sector = 2 if cnae >=10  cnae < 35
replace sector = 3 if cnae >=35  cnae < 41
replace sector = 4 if cnae >=41  cnae < 45
replace sector = 5 if cnae >=45  cnae < 49
replace sector = 6 if cnae >=49  cnae < 55
replace sector = 7 if cnae >=55  cnae < 58
replace sector = 8 if cnae >=58  cnae < 64
replace sector = 9 if cnae >=64  cnae < 68
replace sector = 10 if cnae == 68
replace sector = 11 if cnae >=69  cnae < 77
replace sector = 12 if cnae >=77  cnae < 84
replace sector = 13 if cnae == 85
replace sector = 14 if cnae >=86  cnae < 90
replace sector = 15 if cnae >=90  cnae < 94
replace sector = 16 if cnae >=94  cnae < 97

* Collapse database by survival mean of the aggregated sectors
collapse (mean) s, by(cohort ccaa erte sector t)
* Generating the identifier for any unique observation along time
reshape wide s, i(ccaa cohort erte sector) j(t)
gen id = _n
* Reshaping to long
reshape long s, i(id cohort ccaa erte sector) j(t)
* Setting the panel data structure
xtset id t

* Labels
label var erte "Dummy: Employees in ERTE"
label var t "Quarters since the company was iniatially followed"
label var s "Survival rate regarding the initial cohort"
label var sector "Economic sector of the company"
```

```
label def sector 0 "Total" 1"Extractive industry" 2"Manufacturing" 3"Supply
industry" 4"Construction" 5"Retail" 6"Logistics" 7"Hospitality and tourism"
8"Telecommunications" 9"Finance and insurance" 10"Real estate" 11"Science
and technology" 12"Administrative activities" 13"Education" 14"Health and
social care" 15"Arts and entertainment" 16 "Other services"
label val sector sector
label def ccaa 1 "Andalucía" 2 "Aragón" 3"Asturias" 4"Baleares" 5"Canarias"
6"Cantabria" 7"Castilla y León" 8"Castilla-La Mancha" 9"Cataluña" 10"C. Va-
lenciana" 11"Extremadura" 12"Galicia" 13"C. Madrid" 14"R. Murcia" 15"C.F.
Navarra" 16"País Vasco" 17"La Rioja" 18"Ceuta" 19"Melilla" 0"España"
label val ccaa ccaa

* Dropping totals
keep if cohort==1  ccaa<18
drop if erte==3 | ccaa==0 | sector==0
compress
save "data_mainsectors.dta", replace // saving database


***********************************************************
*** DATABASE 2: BY REGIONS, EMPLOYEES AND ERTE ***
***********************************************************
clear all
set more off
capture log close
cd "C:WorkingDirectory" // Set your working directory

** APRIL 2020 COHORT DATA **
* Setting initial values of global macros for the loop
global inicio 6
global final 24
global i 0
while $i < 20 {
   import excel "C:WorkingDirectoryAndFileName.xlsx", sheet("1 ABRIL 2020")
   cellrange(A$inicio :K$final ) firstrow case(lower) // import code
   * format of survival rates *
   local S s1 s2 s3 s4
   foreach s in `S' {
      destring `s', float replace
      replace `s' = `s'/100 // survival in fraction instead of percentage
      }
   * generating regional datasets *
   reshape long q s, i(emp erte) j(t) // reshape to long data
   drop if erte == 3 // dropping totals (sum of erte and non erte obs)
```

```
    gen cohort = 1 // id for cohort
    gen ccaa = 0 + $i // id for region
    replace s = 1 if t==0 // survival rate in the initial period == 1 always
    save "edata_cohort1_region$i.dta", replace
    clear
    * iterating the loop *
    global i = $i + 1
    global inicio = $inicio + 22
    global final = $final + 22
    }

* Appending regional datasets
use "edata_cohort1_region0.dta", clear // master dta file
forval n = 1/19 {
    append using "edata_cohort1_region`n'.dta" // append loop
    erase "edata_cohort1_region`n'.dta" // erasing unnecessary files
    }
save, replace // saving database

* Labeling variables
label var emp "Number of employees in the company (categorical)"
label var erte "Dummy: Employees in Temporal Employment Regulation situation (ERTE)"
label var t "Quarters since the company was iniatially followed"
label var q "Number of surviving companies"
label var s "Survival rate regarding the initial cohort"
label def emp 0 "Total" 1 "From 1 to 5" 2 "From 6 to 9" 3 "From 10 to 99" 4 "From 100 to 249" 5 "More than 250"
label val emp emp
label def ccaa 1 "Andalucía" 2 "Aragón" 3"Asturias" 4"Baleares" 5"Canarias" 6"Cantabria" 7"Castilla y León" 8"Castilla-La Mancha" 9"Cataluña" 10"C. Valenciana" 11"Extremadura" 12"Galicia" 13"C. Madrid" 14"R. Murcia" 15"C.F. Navarra" 16"País Vasco" 17"La Rioja" 18"Ceuta" 19"Melilla" 0"España"
label val ccaa ccaa

* Generating the identifier for any unique observation along time
sort cohort ccaa emp erte t
egen id = group ( cohort ccaa emp erte )
* Setting the panel data structure
xtset id t
order id cohort ccaa t erte emp q s
compress
save "data_employees.dta", replace // saving database
erase "edata_cohort1_region0.dta"
```

```
**********************************************************

********************
** DESCRIPTIVES **
********************
clear all
capture log close
cd "C:WorkingDirectory" // Set your working directory


** SECTORs data **
frame create sectors
frame change sectors
use "data_sectors.dta"


* Dropping initial observations for the analysis
drop if t==0


* Descriptive tables
sum s ,detail
eststo clear
estpost tabstat s , stats(mean sd min max p25 p50 p75 count )
esttab . using descriptives.tex , cells("mean sd min max p25 p50 p75 count")
label noobs title("Survival rate descriptive stats") notes("Data source: CO-
DEM.")
tabstat s , stats(mean sd min max p25 p50 p75 count ) by(erte)
tabstat s , stats(mean sd min max p25 p50 p75 count ) by(ccaa)
tabstat s , stats(mean sd min max p25 p50 p75 count ) by(t)
tabstat s , stats(mean sd min max p25 p50 p75 count ) by(sector)


* Descriptive graphs
graph hbox s, over(ccaa, sort(1) des) name(by_ccaa1)
graph hbox s, over(sector, sort(1) des) name(by_sector1)
graph hbox s, over(erte, sort(1) des) name(by_erte1)
graph hbox s, over(t, sort(1) des) name(by_t1)


** SIZE-Employees data **
frame create size
frame change size
use "data_employees.dta"


* dropping initial observations
```

```
drop if t==0

* Descriptive tables
estpost tabstat s , stats(mean sd min max p25 p50 p75 count )
esttab .  using descriptives.tex , cells("mean sd min max p25 p50 p75 count")
label noobs title("Survival rate descriptive stats") append
estpost tabstat s , stats(mean sd min max p25 p50 p75 count ) by(erte)
estpost tabstat s , stats(mean sd min max p25 p50 p75 count ) by(ccaa)
estpost tabstat s , stats(mean sd min max p25 p50 p75 count ) by(t)
estpost tabstat s , stats(mean sd min max p25 p50 p75 count ) by(emp)

* Descriptive graphs
graph hbox s, over(ccaa, sort(1) des) name(by_ccaa2)
graph hbox s, over(emp, sort(1) des) name(by_emp2)
graph hbox s, over(erte, sort(1) des) name(by_erte2)
graph hbox s, over(t, sort(1) des) name(by_t2)

* Combined graphs
graph combine by_ccaa1 by_ccaa2 , col(2) row(1) name(by_ccaa)
graph combine by_erte1 by_erte2 , col(2) row(1) name(by_erte)
graph combine by_t1 by_t2 , col(2) row(1) name(by_t)

*************************************
**** FRACTIONAL REGRESSIONS *****
*************************************

* SECTOR DATASET *
frame change sectors

eststo clear
fracreg logit s erte ib1.t ib1.ccaa ib7.sector

* Marginal effects
eststo margin1: margins, dydx(*) post
margins, dydx(erte) at(t=(1 (1) 4))
marginsplot, name(marginplot1)
margins, dydx(erte) at(sector=(1 (1) 16))
marginsplot

* Note: dy/dx for factor levels is the discrete change from the base level.

***************************************

* SIZE DATASET *
frame change size
```

```
fracreg logit s erte ib1.t ib1.ccaa ib1.emp

* Marginal effects
eststo margin2: margins, dydx(*) post
margins, dydx(erte) at(t=(1 (1) 4))
marginsplot, name(marginplot2)
margins, dydx(erte) at(emp=(1 (1) 5))
marginsplot

* Export margins table and graphs
esttab using margin.tex, b(4) p nostar wide nobaselevels title(Average marginal
effects on survival rate.) label addnotes(Baseline categories: ERTE=0, 1-5 em-
ployees, Trend dummy=1, Hospitality and tourism, Andalucía.) mtitles(Sector_database
Size_database) replace
graph combine marginplot1 marginplot2 , col(1) row(2)
```