Rise and fall of the largest firms in Portugal

João Amador
Banco de Portugal
Nova School of Business and Economics

Mário Lourenço
Banco de Portugal

Cloé Magalhães
Banco de Portugal

Ana Catarina Pimenta
Banco de Portugal

October 2020

Abstract
Several studies have been establishing stylized facts regarding firm entry and exit of the market. However, there is still limited evidence on the dynamics of top firms throughout very long periods of time. This article builds on a new database that records the largest firms operating in Portugal in terms of annual turnover. We consider the set of the top 200 firms in the period 1981-2018 and assess their dynamics across different classes in the ranking and the probability of exiting this group. The article concludes that there is more stability in terms of the firms placed in the top classes of the ranking. Moreover, on average, for different ranking classes and time horizons, the probability of rising in the rank is smaller than the probability to fall. The survival in the ranking differs according to the sector in which firms operate. Firms in the electricity and water supply sector survive for longer periods, while the median duration is lower in industry and construction. (JEL: L11, L20, L25)

"All live to die, and rise to fall."
Christopher Marlowe

1. Introduction

Economic literature refers firms' demography as a driver of economic growth. One relevant dimension is the entry and exit of firms in the market – the designated extensive margin. The relationship with economic growth is linked with the idea of "creative destruction" by Schumpeter (1911, 1942). According to this view, firms that enter the market bring new goods and services that, if successful, will replace outdated ones. This process makes firms that produce outdated goods or services exit the market, generating short-term losses in activity that will be more than offset in the medium and long term, thus bringing net gains in value added. For new firms whose goods

Acknowledgements: The authors thank Beatriz Pires and Bernardo Sarrasqueiro for the excellent work on the compilation of the dataset. The authors are also thankful to Nuno Alves, António Antunes, Carlos Gouveia, Filipa Lima and Pedro Duarte Neves for their useful comments and suggestions. The analyses, opinions and conclusions expressed herein are the sole responsibility of the authors and do not necessarily reflect the opinions of Banco de Portugal or the Eurosystem.

E-mail: jamador@bportugal.pt; mfllourenco@bportugal.pt; clmagalhaes@bportugal.pt; apimenta@bportugal.pt

or services do not pass the test of market, there should be a swift exit process, freeing up resources for new start-ups or for the growth of incumbents.

Another important dimension of firms' demography is their growth – the designated intensive margin – measured in terms of turnover, gross value added or employment. Firms' growth dynamics depend on multiple factors. Some of them relate to the specific characteristics of the firm, such as its orientation towards foreign markets, the ability to innovate and differentiate its goods or services or the quality of management. Other factors are external to the firm and relate with the regulatory environment, financing conditions, inefficiencies in the functioning of labour and product markets and overall macroeconomic developments. These elements also have a bearing on the extensive margin. For example, the underlying market competition conditions influence both firms' entry and exit, as well as changes in their market shares.

The empirical literature on firms' demography is too large to be fully mentioned here. Given their importance for long-term economic growth, much of the literature focuses on startups' probability of survival. Some contributions are those of Wagner (1994), Audretsch *et al.* (1999), and Mata *et al.* (1995), which analysed the post-entry performance of new manufacturing firms in Italy, Germany and Portugal, respectively. In a similar strand of work, several studies have pointed out that startups' survival depends on initial size (e.g. Mata and Portugal (1994) and Mata *et al.* (1995)), age (e.g. Dunne *et al.* (1989)), bank funding (e.g. Farinha *et al.* (2019)), human capital in the firm (e.g. Mata and Portugal (2002)), among others. A totally different strand of literature focuses on the role of very large firms and their granular impacts on aggregate outcomes. This literature started with the seminal work by Gabaix (2011) and a related application using data for Portuguese exporters is that of Cabral *et al.* (2020). A third strand of literature focuses on the behaviour of very large firms and the rise in product market concentration. Examples of this literature are Autor *et al.* (2017), which associates superstar firms to the fall in labour share in the US, and OECD (2018), which surveys information on market concentration in the US, Japan and Europe. Nevertheless, our study does not directly relate to these strands of analysis.

Possibly due to lack of information, the literature has not been covering the evolution of firms over long periods of time, which limits the ability to fully assess their rise-fall dynamics. Information about the year of creation of a firm allows for the estimation of survival functions that convey information on the probability of survival in each moment, but this is different from assessing the individual path of each firm regarding its relative size in the economy. Nevertheless, although existing studies cover relatively short time spans, they consider large sets of firms, which increases the representativeness of results. These aspects determine the unavailability of analyses that could be deemed comparable to our work and, hence, used to benchmark its results.

In this article we contribute to the literature by considering firms' rise-fall dynamics in wider time horizons. We use a database that identifies the largest Portuguese firms, which spans over four decades and was built for this purpose from previously scattered business information. More specifically, the sample used in this analysis covers the top 200 firms in terms of turnover in each year between 1981 and 2018. The methodological approach is mostly descriptive, i.e., we are not suggesting explanatory factors for

the identified firms' rise and fall dynamics. Results point to a higher stability of the firms in the top classes of the ranking when compared with those in lower classes, thus indicating that the largest firms in the ranking generally maintain their positions. Moreover, on average, the probability of rising in the ranking tends to be smaller than the probability of falling, which is explained by the effect of firms entering directly to intermediate positions in the ranking.

The article is organized as follows. The second section presents the details of the database. Additionally, descriptive statistics regarding the sectoral dimension and the distribution of the largest 200 firms are briefly analysed. The third section presents the results of the rise and fall dynamics of firms in the top 200 during the period considered. Firstly, we present the transition matrices between ranking classes for different time horizons. Secondly, survival functions are estimated, relating the number of years in the ranking and the probability of staying in. The fourth section includes some concluding remarks.

2. Database and descriptive statistics

This section presents the data sources and procedures underlying the construction of the *Largest Portuguese Firms Database*. This dataset was constructed with a view to support this article but it will be made available for further research. In addition, this section provides a brief set of descriptive statistics regarding the referred dataset.

2.1. Data sources, treatment and harmonization

The *Largest Portuguese Firms Database* contains the top Portuguese firms in terms of turnover, mainly combining data from the Simplified Corporate Information (IES, the Portuguese acronym for "*Informação Empresarial Simplificada*"), which contains information on balance sheets and income statements for almost all non-financial Portuguese firms from 2006 onwards, and for more remote periods, data directly collected from hard copies of specialized business publications that summarize publicly available information. The *Largest Portuguese Firms Database* covers the 1976-2018 period, including each firm's yearly turnover (ranking variable).

For the years 1992 onwards, the database uses the reference population of active Portuguese firms estimated by the Statistics Department of Banco de Portugal. This dataset contains, for each firm, variables such as the identification number, headquarters location and the main sector of economic activity according to NACE Rev.2, as well as the number of employees, turnover, total assets and capital. This reference population collects information from several sources. Besides the already mentioned major contribution from the IES database, the Central Registry of Legal Entities, a database managed by the Institute of Registries and Notary of the Ministry of Justice, the Statistical Units File of Statistics Portugal, the Quarterly Survey on Non-Financial Corporations (ITENF, the Portuguese acronym for "*Inquérito Trimestral às Empresas Não Financeiras*"), the Integrated System of Securities Statistics, Banco de Portugal's Central

Credit Register and information obtained for the compilation of the Portuguese Balance of Payments and International Investment Position statistics are all taken into account.

Up to 1992, the yearly rankings of Portugal's largest firms published by specialized business publications were used to feed the database. Several publications were combined in order to obtain the largest information set possible. For the years from 1976 to 1978, data was collected from the *SEMAP* ranking of Portuguese top firms; for the years between 1979 and 1990, such ranking was collected from the *Negócios* magazine and, for 1991, the *Exame* magazine was considered. For each of these sources, all data deemed relevant was manually inserted into the database: company name, number of employees in each year, turnover, among other elements. Nevertheless, the number of top firms was not consistent across publications: the rankings ranged from the top 100 companies (for the most remote years) to rankings containing the top 500 companies in each year.

As the *Largest Portuguese Firms Database* combines several sources of information, some harmonization procedures were required. Given the need to uniquely identify each company throughout the relevant time span of the database, the most recent firm identification number was used. The link between the firm's most recent identifier and previous ones was manually established whenever possible. When this procedure was not possible (for example, in the cases of firms included in remote years that have ceased activity or that have merged with new firms) specific codes where attributed to identify the firms throughout the entire database. Firms' classification by economic activity, institutional sector and location of their headquarters were considered the same throughout the time span of the database, corresponding to the most recent information available.

Most firms included in the *Largest Portuguese Firms Database* belong to the Non-Financial Corporations sector, as defined by the European System of Accounts (ESA 2010). Given their importance in some activities, state owned firms are also included, belonging to the Non-Financial Corporations or the General Government sectors. Non-financial holdings which are classified as Financial Corporations are also considered in the database in order to cover the activity of economic groups.

A database comprising only a small set of large firms is necessarily affected by undesired attrition. This can occur due to the creation or exit of large special purpose entities that may not have a connection with actual economic activity or due to mergers and acquisitions. These problems were addressed in several ways. Firstly, firms established in the Madeira Free Trade Zone, whose participation in the ranking was relevant for several years, were removed.[1] Secondly, specific events involving the top ranked firms (e.g. mergers, spin-offs, etc.) were also addressed. These events may have resulted solely from the restructuring of groups, hence leading to artificial cases of entry

1. The Madeira Free Trade Zone (*Zona Franca da Madeira*) is in activity since 1986. Several large special purpose entities established headquarters in this location with a view to benefit from special tax conditions. The original database included 82 entities located in the Madeira Free Trade Zone, most of which recorded in the 1995-2003 period and in the "Transportation and communication" and "Other activities" sectors. These entities were removed from the final version of the database.

and exit from the database. In such cases, whenever possible, the entities involved were aggregated and a single identification number was considered throughout the time span of the *Largest Portuguese Firms Database*. These adjustments made it possible to take into consideration the restructuring of large economic groups, even if the dissolution of some holding companies could be corrected.

For the purpose of this article, the top 200 firms listed in the *Largest Portuguese Firms Database* were considered from 1981 to 2018. Data from 1976 to 1980 were excluded because the database lists less than 200 companies in those years. Overall, our final selection contains 7600 observations, comprising 835 distinct firms and 38 years of information.

2.2. Descriptive statistics

The share of the top 200 firms on total Gross Value Added (GVA) of the Portuguese economy is quite high. These firms represented about 10% of Portugal's total GVA in 2018 (Figure 1). This share has decreased since 1995, when it represented approximately 15% of total GVA. This evolution was mostly determined by the decrease of the share of the GVA generated by the top 50 firms of the ranking, in particular by firms in the "Electricity and water" and "Transportation and communication" sectors. The share of these firms relatively to the respective sectoral GVA is higher than in the remaining sectors, having decreased from close to 96% and 83%, respectively in 1995 to values close to 40% in 2018. As for the remaining activities, firms in "Industry", "Construction" and "Trade, accommodation and food services" in the ranking present a stable share of the respective sectoral GVA, fluctuating between 5% and 20% in the horizon under analysis.

Considering the set of firms included in the rankings at least once during the 1981-2018 period, 35% belong to the sector "Trade, accommodation and food services" and 31% to "Industry". "Construction" and "Transportation and communication" represent 10% and 7% of firms in the ranking, respectively, while "Electricity and water" represents 2% of firms. Those operating in other activities correspond to 15% of the firms in the ranking.

The distribution of the top 200 firms by sector in each year is shown in Figure 2. "Industry" represented 37% of this set of firms in 1981, reaching a maximum of 47% in 1985. In the two decades that followed, the share of "Industry" in the set of companies analysed decreased to a minimum of 19% in 2005. After the global economic and financial crisis of 2008, the relevance of this sector steadily increased, standing at 32% of the set of largest companies analysed in 2018.

The sector "Trade, accommodation and food services" has shown an inverse trend. The share of firms in the top 200 operating in this sector increased from 27% in 1985 to 48% in 2005. Later, this proportion dropped to around 40%. As for the remaining activities, "Construction" represented around 5% of the firms in the ranking until 1994. In the two decades that followed, the share of this sector in the ranking stood at around 8%. After reaching a maximum of 11% of the firms in the ranking in 2011, the share of "Construction" decreased to a minimum of 3% in 2018.

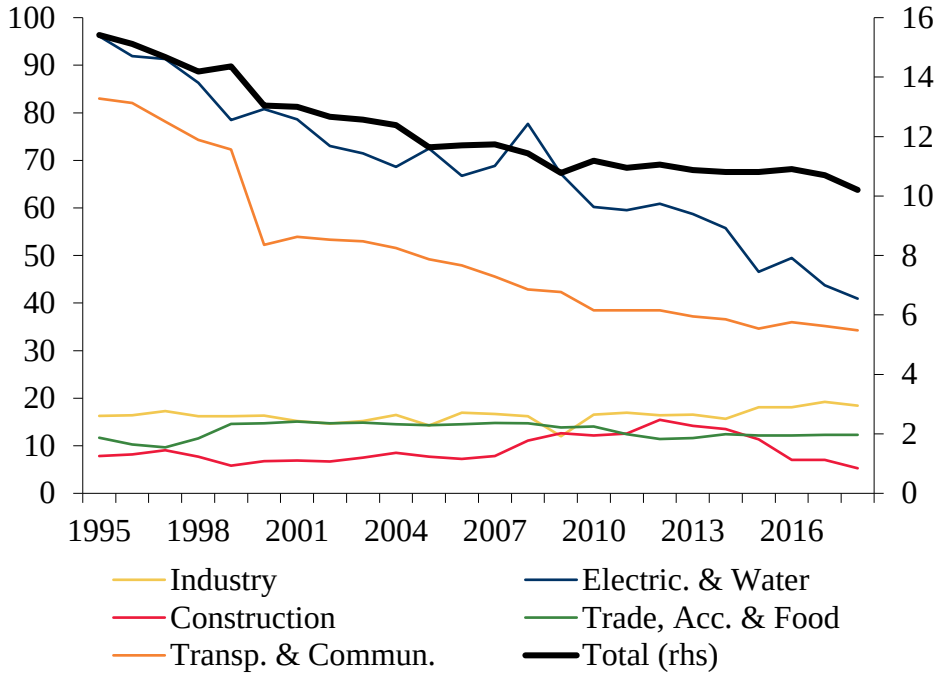


FIGURE 1: Share of top 200 Portuguese firms on Gross Value Added (GVA), percentage

Source: Banco de Portugal and Statistics Portugal.
Notes: Gross Value Added of the economy available as of 1995. The sectors presented in this figure correspond to aggregations of NACE Rev.2 sections: "Industry" (sections B and C), "Electricity and water" (sections D and E), "Construction" (section F), "Trade, accommodation and food services" (sections G and I), "Transportation and communication" (sections H and J). The GVA generated by the "Transportation and communication" firms in the top 200 is affected by the restructuring of a "Communications" group in 2000.

Finally, "Electricity and water" increased its relevance in the rankings, from around 1% in the early 1980s to 6% in 2018. This evolution reflects the developments in the electricity market during the last decades, notably the privatization of *Energias de Portugal* in the late 1990s and the segmentation of production, distribution and retail activities imposed by the implementation of the Iberian Energy Market as of 2006. The restructuring of these companies led to the establishment of new operators that have become a part of the list of the top 200 largest Portuguese firms.

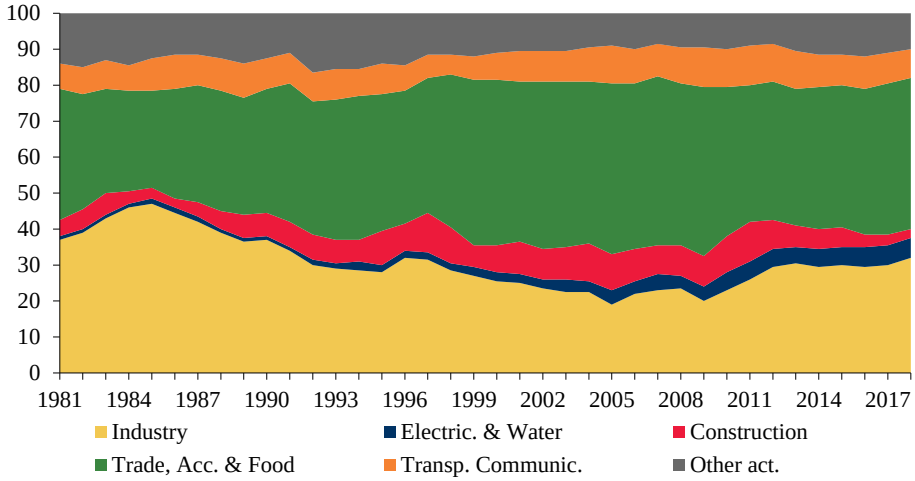
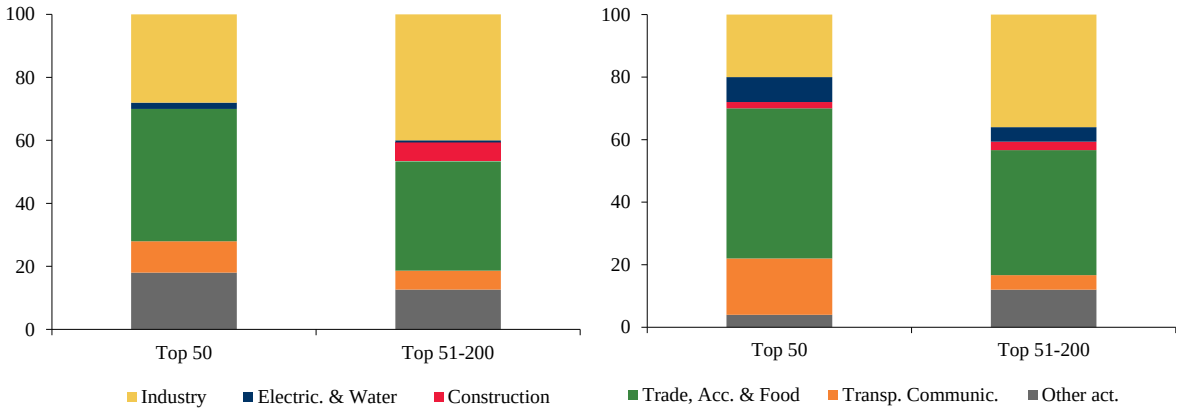


FIGURE 2: Distribution of the top 200 Portuguese firms by sector and year

Notes: The sectors presented in this figure correspond to aggregations of NACE Rev.2 sections: "Industry" (sections B and C), "Electricity and water" (sections D and E), "Construction" (section F), "Trade, accommodation and food services" (sections G and I), "Transportation and communication" (sections H and J).

The sectoral structure of the top 50 firms also provides further information (Figure 3). In 2018, "Trade, accommodation and food services" stood for nearly half of the top 50 largest firms, while representing 40% of the remaining top 51-200 firms. The relevance of network industries among the top 50 largest firms is also noticeable. "Transportation and communication" and "Electricity and water" sectors were more relevant among the top 50 largest firms (18% and 8%, respectively) than among the remaining top 51-200 firms (5% in both sectors). Conversely, firms in the "Industry" sector represented 20% of the top 50 largest firms and 36% of the remaining top 200 firms. "Industry" firms in the top 50 mostly operate in fuel, transport equipment and components, as well as food and drinks industries, while the activities of "Industry" firms in the remaining top 51-200 are more disperse. The structural differences between the top 50 largest firms and the remaining top 51-200 are observable throughout the time span covered, as shown when comparing the situation in 1981 and 2018.

Another dimension of analysis is the number of years that each firm belongs to the top 200. The average number is 9 years and the distribution is highly skewed to the right, which means there is a large number of firms that take part of the ranking for a relatively small number of years (Figure 4). The average is quite misleading to describe the individual developments of firms in the economy (Altomonte *et al.* 2011). The median firm was part of the top 200 for 6 years; on the other hand, 6% of the firms (46 out of 835 firms) were part of the top 200 for at least 30 years.



(A) 1981 (B) 2018

FIGURE 3: Distribution of the top 200 Portuguese firms by position within the ranking (top 50 vs top 51-200) and sector

Notes: The sectors presented in this figure correspond to aggregations of NACE Rev.2 sections: "Industry" (sections B and C), "Electricity and water" (sections D and E), "Construction" (section F), "Trade, accommodation and food services" (sections G and I), "Transportation and communication" (sections H and J).

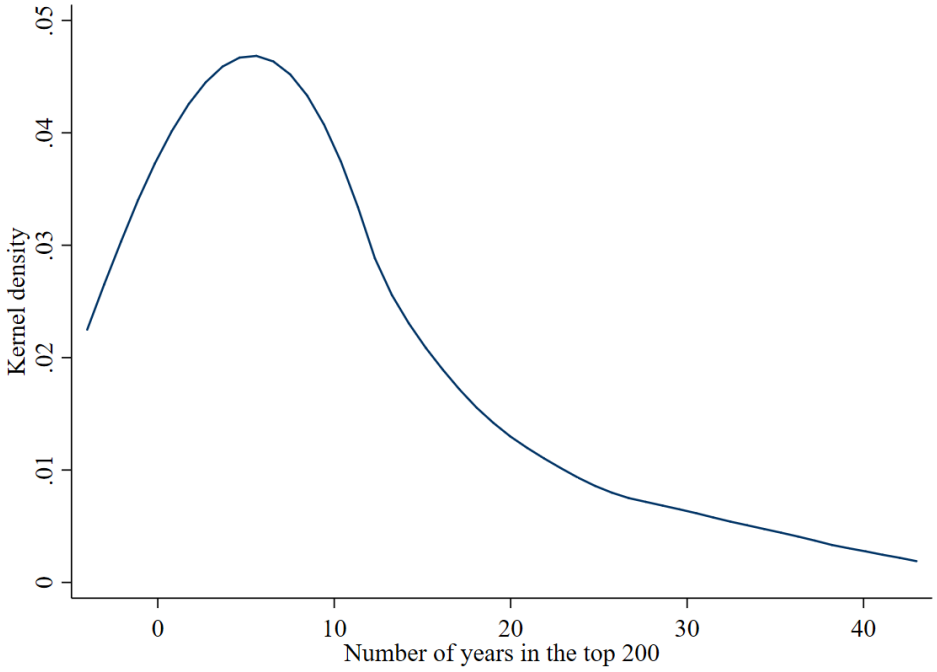
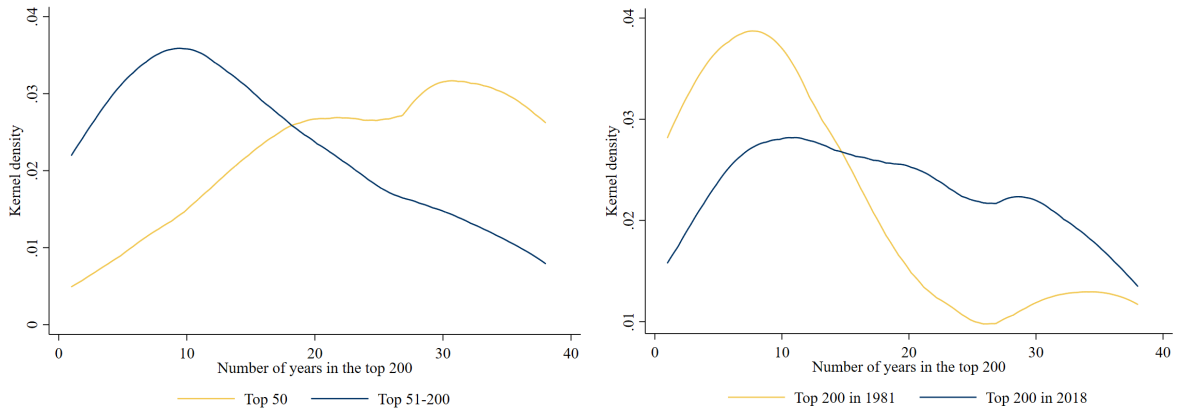


FIGURE 4: Distribution of the number of years in the top 200

Notes: The number of years in the top 200 was uniquely computed for each of the 835 distinct firms that were part of it between 1981 and 2018. The Kernel density estimation is a nonparametric method of estimating the probability density function of a variable. In the literature, for continuous variables, these density estimates are considered preferable to histograms because they smooth the distribution.

As previously mentioned, the top 50 firms show a different profile than the rest of those in the top 200 (Figure 5, panel A). As expected, it is noticeable that the former firms stay in the top 200 longer than the rest. In the 38 years covered in this analysis,

firms in the top 50 were a part of the ranking, on average, 17 years (median of 18 years), while the remaining ones have been part of the top, on average, almost 8 years (median of 5 years). Furthermore, firms that were part of the top 200 in the later years of the sample remain there for longer periods than those identified in the earlier years (Figure 5, panel B). The median number of years in the ranking increased from 10, for the top 200 firms in 1981, to 18, for the top 200 firms in 2018, thus pointing towards a higher stability in the ranking in the last decades. These features are further developed in the next two sections by analysing transition matrices between classes of positions in the ranking and by estimating survival functions.



(A) By class of positions (B) By year

FIGURE 5: Distribution of the number of years in the ranking

Notes: The number of years in the top 200 was uniquely computed for each of the 835 distinct firms that were part of the top 200 between 1981 and 2018. In Panel A, the "Top 50" correspond to those firms whose most frequent position in the ranking is in the top 50 positions. "Top 51-200" are the remaining firms. Panel B describes the distribution of the number of years in the top for those firms that were observed in the 2018 and 1981 top 200 firms.

3. Results

The dynamics in the ranking of the 200 largest firms according to turnover is analysed along two complementary approaches. Firstly, transition matrices for different time spans inform on the conditional probability of firms moving across pre-established ranking intervals. Secondly, considering only firms whose entrance in the ranking is identified in a given moment in time, estimated survival functions inform on the likelihood of them remaining in the ranking within different time spans.

3.1. Transition matrices

The results presented in this subsection are based on eight classes of positions in the ranking of the 200 largest firms according to turnover, as described in subsection 2.1. The eight classes considered correspond to positions: [1-25]; [26-50]; [51-75]; [76-100]; [101-125]; [126-150]; [151-175]; [176-200], to which is added the status "Out" of the ranking. In each year, the firms that belong to the "Out" category correspond to those that have at least once been included in the ranking but do not belong to the ranking in that specific year.

The initial analysis assesses the overall dynamics of firms among classes. The shape of the distribution of the number of changes in class of positions for each firm, considering different time spans of permanence in the ranking, is presented in Figure 6. As expected, the median number of changes increases for larger time spans. The median number of changes between classes is 4 for firms that belong to the ranking between 6 and 10 years and increases to 11 and 9 for those that belong to the ranking between 26 and 30 years and between 31 and 38 years, respectively. Moreover, the median and the percentile 25 are lower for those firms that belong to the ranking between 31 and 38 years comparing to those that remain in the ranking between 26 and 30 years. This points towards more stability when firms are a part of the ranking for longer periods.

The previous analysis can be developed by explicitly taking into account the movements observed between specific classes. Benefiting from the long time horizon available in the database, transition matrices are consecutively computed for intervals between 1 and 20 years. As an illustration, Table A.2 in Appendix represents the transition matrix for a time horizon of 10 years. Rows identify the starting position of the firm in moment t and columns refer to its position in the period $t + 10$, thus, each row is a conditional distribution adding up to 100%. According to Table A.2, the 200 largest firms tend to remain in the same ranking class 10 years later, i.e., the probabilities in the main diagonal of the transition matrix are higher, particularly in the upper classes. For example, 48.4% of firms in the interval [1-25] in a given year remain within that same set of positions 10 years later. Other firms fall within the ranking: 7.4% move to the interval [51-75] and 23.4% exit the ranking. As anticipated, an important feature in the transition matrix is that firms in the lower intervals have a lower probability of remaining in the ranking. For example, only 4.3% of firms in the interval [176-200] remain in that class 10 years later, while 80.1% of them exit the ranking.

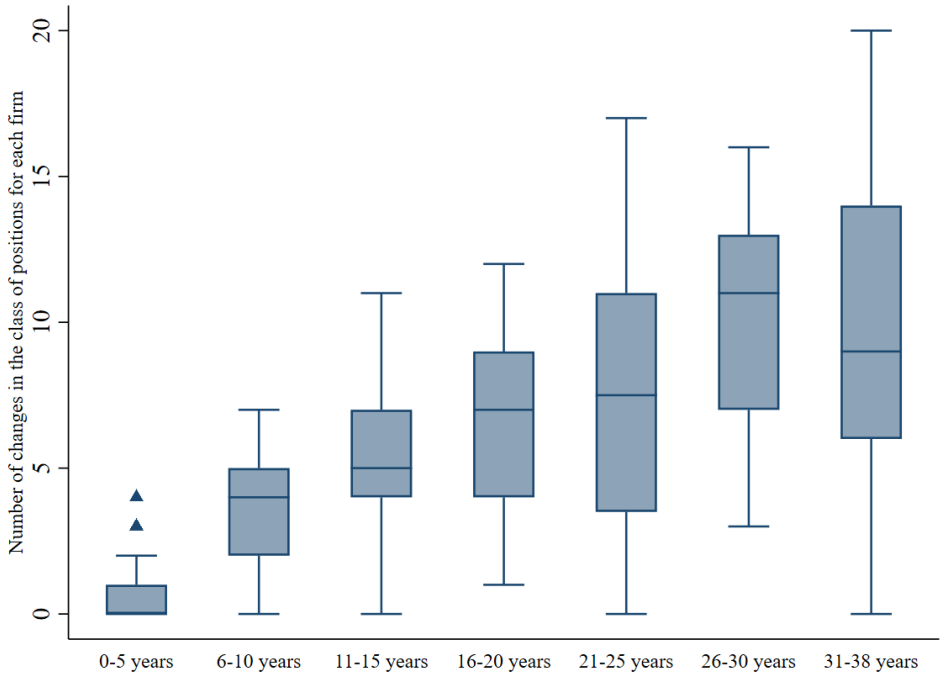


FIGURE 6: Distribution of changes between classes of positions for each firm, by the number of years in the top 200

Notes: In each year, it is assumed that a firm changes in the class of positions if the class to which it belongs in that year is different from the class to which it belonged in the previous year. If the firm belongs to the ranking for only one year, there are no changes. The classes considered correspond to positions: [1-25]; [26-50]; [51-75]; [76-100]; [101-125]; [126-150]; [151-175]; [176-200], to which is added the status "Out" of the ranking. In the box plot, the central box represents the values of the 25th percentile to 75th percentile (interquartile range) and the horizontal line correspondes to the median of the distribution (50th percentile). The vertical line extends from the minimum to the maximum value, excluding outliers (values lower than the difference between the 25th percentile and 1.5 times the interquartile range, or higher than the sum of the 75th percentile and 1.5 times the interquartile range). The triangles correspond to outliers.

The information contained in the transition matrices for different time spans allows for the identification of other stylized facts. Figure 7 presents the probability of firms placed in each class of the ranking to remain in the same class of positions up to 20 years later.[2] As illustrated in Appendix A, these "post entry" probabilities correspond to the cells in the main diagonal for the successive transition matrices from 1 up to 20 years. For example, results indicate that firms in the class [1-25] have a slowly decreasing probability of keeping this position in the future, standing at 87% and 81% after 1 and 2 years, respectively, and standing at around 29% after 20 years. For the firms placed in the class [26-50] the probabilities of keeping their position also decrease as the time span widens, reaching approximately 17% after 20 years, while in the lower classes of the ranking the probabilities decay faster and stand below 5% after 20 years. Overall, one important result that emerges is the relative resilience of the top firms in the ranking. The Spearman's rank correlation coefficient is an alternative measure of the stability of firms between classes of positions in the ranking. On average, the correlation coefficient

2. Longer transitions could be considered but the number of underlying firms used in the computation would be smaller and results would thus be less robust.

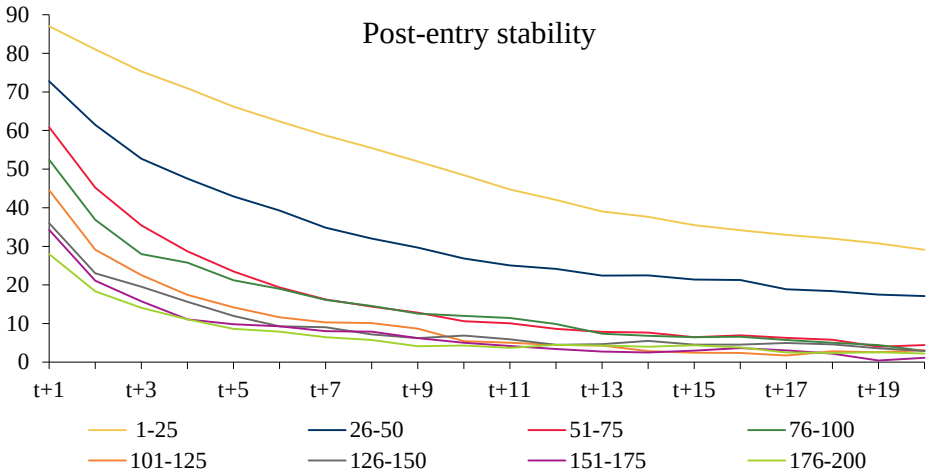


FIGURE 7: Probability of remaining in the same class of positions in the ranking between year t and year t+x, with x varying between 1 and 20 years

Notes: For each line, values correspond to those of the main diagonal of each of the 20 transition matrices. For instance, for the top 25 firms, in each year, the probability of remaining in that category 1 year later corresponds to the first cell of Table A.1 (87%). Similarly, the probability of these firms remaining in the top 25 after 10 and 20 years corresponds to the first cell of Table A.2 (48.4%) and Table A.3 (29.1%), respectively.

between the firms' class in year t and year $t + x$ tends to be lower as x widens, thus confirming the relative stability for shorter horizons (Figure B.1 in appendix).

Another perspective is to assess the probability of firms moving to a higher class in the ranking in periods from $t + 1$ up to $t + 20$, depending on the class where they start. The probability of firms "rising" in the ranking corresponds, in each line of the transition matrix, to the horizontal sum of cells to the left of the main diagonal, along the different time horizons. Results are presented in Figure 8. The probability of rising in period $t + 1$ coming from the class [26-50] is about 9.7%, it increases to 13.3% nine years later and drops to 5.1% after 20 years. Conversely, the probability to rise starting from the lowest ranking class [176-200] in period $t + 1$ is 25.6%, decreases to 15.6% in $t + 10$ and declines until 9.3% in $t + 20$. Therefore, as expected, it is easier to rise when starting from below but this feature is not as strong in larger horizons.

The dynamics of firms falling or exiting from the ranking is described in Figure 9. In this case, the probability of fall or exit from the ranking in time horizons from t up to $t + 20$, when starting from each class, is equivalent to the horizontal sum of the cells to the right of the main diagonal (i.e., including the "Out" category) for each line.[3] Results show that the probability of fall or exit from the ranking by firms in interval [1-25] is 13% in period $t + 1$, 51.6% in $t + 10$ and 70.9% in $t + 20$. Conversely, the probability of exiting for firms in the interval [176-200] (falling is not possible) is 46.4% in period $t + 1$, 80.1% in $t + 10$ and 88.4% in $t + 20$.

At this point it is relevant to highlight that, for each starting class and different time transitions, the probability of falling or exiting the ranking is larger than that of rising,

3. It is worth noting that, for each class of positions, the sum of the probabilities of stability (Figure 7), rise (Figure 8) and fall or exit (Figure 9) corresponds to 100%.

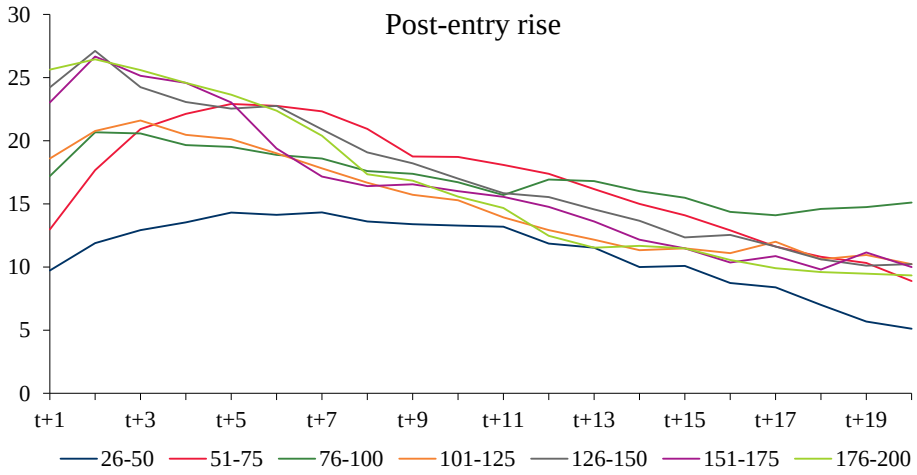


FIGURE 8: Probability of rising in the ranking in year $t + x$, given that in year t the firm belongs to each of the indicated class of positions, with x varying between 1 and 20 years

Notes: For each line, values correspond to the horizontal sum of the transition probabilities to the left of the main diagonal. For instance, considering that in year t the firm belonged to the top 26 to 50, the probability of rising to the top 25 is 9.7% after 1 year (Table A.1), 13.3% after 10 years (A.2) and 5.1% after 20 years (A.3). Similarly, for firms in the top 51 to 75 in year t, the probability of rising in the ranking is 13% after 1 year (Table A.1), 18.7% after 10 years (Table A.2) and 8.9% after 20 years (Table A.3). Given the classes of positions considered, it is not possible to rise when firms are already in the top 25.

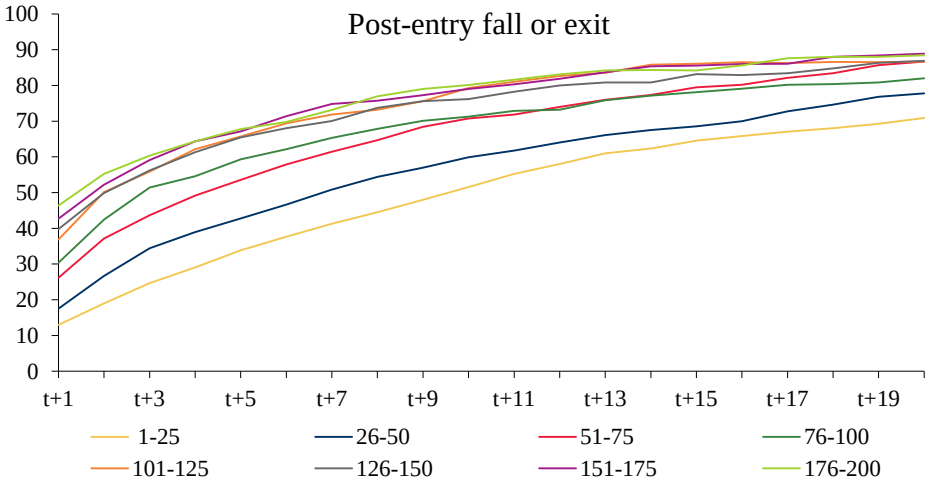


FIGURE 9: Probability of falling or exiting the ranking in year $t + x$, given that in year t the firm belonged to each class of positions, with x varying between 1 and 20 years

Notes: For each class of positions in the ranking, the values in the lines correspond to the sum of the transition probabilities to the right of the main diagonal. In each year, the firms that belong to the "Out" category correspond to those that have at least once been included in the ranking of the 200 largest firms, but do not belong to the ranking in that year. For instance, considering that in year t the firm belonged to the top 25, the probability of fall or exit from the ranking 1 year later is 13% (Table A.1) and 51.6% and 70.9% after 10 and 20 years, respectively (Table A.2 and Table A.3, respectively).

i.e., in each line, the sum of cells to the left of the main diagonal is smaller than the sum of cells to the right. This strong regularity is an important result and has a bearing on the perception about the dynamics of the largest firms in the market. Even if reaching the ranking signals success, the sword of Damocles is always hanging over their head.

Complementary, we focus on the path of firms entering the ranking. The probability of moving to the different intervals in the ranking in periods $t+1$ up to $t+20$ when starting from the situation "Out" corresponds to the bottom line in the different transition matrices, as signaled in Appendix A. Results are represented in Figure 10 and each line identifies the probability of an outside firm to move to the corresponding ranking interval in each time horizon. The probability of moving to each interval is smaller the higher the classes in the ranking. Moreover, the probability of ascending to each class increases along time. Nevertheless, it should be noted that the probabilities are relatively low in all horizons.

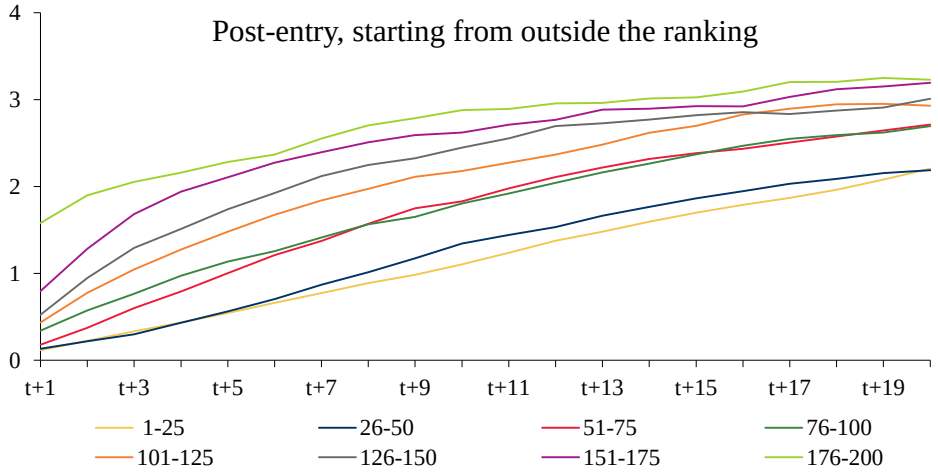


FIGURE 10: Probability of belonging to each class of positions in year t+x, given that in year t the firm was out of the ranking, with x varying between 1 and 20 years

Notes: For each of the 20 transition time horizons considered, the values correspond to those of the line highlighted in yellow (see Tables A.1, A.2 and A.3 as examples). In each year, the firms that belong to the "Out" category correspond to those that have at least once been included in the ranking of the 200 largest firms, but do not belong to the ranking in that year. In each year, 0.1% of the firms that were out of the ranking transitioned to the top 25 and 1.6% to the top 176 to 200 after 1 year (Table A.1). These probabilities are 1.1% and 2.9% for a 10-year transition horizon (Table A.2) and 2.2% and 3.2% for a 20-year transition horizon (Table A.3).

Finally, the dynamics of firms before exiting the ranking ("pre-exit") is described in Figure 11. In this case, taking firms that exit the ranking in moment t we analyse the probability of those firms being in each interval in periods from $t-1$ up to $t-20$. This information corresponds to the latest column in the set of the successive transition matrices, as signaled in Appendix A. Results show that the probability of exit from the ranking by firms that in the year before were in classes [1-25] and [26-50] is smaller (2.6% and 2.4%, respectively) increasing up 39.8% and 48.9% when the time horizon recedes to 20 years before exit. Conversely, firms present in the class [176-200] in the moment prior to exit have a probability of 46.4% of exiting, which increases to 88.4% if they depart from this same class 20 years earlier. This confirms the result of stability in the higher positions of the ranking, i.e. larger firms have a relatively higher probability of maintaining their positions.

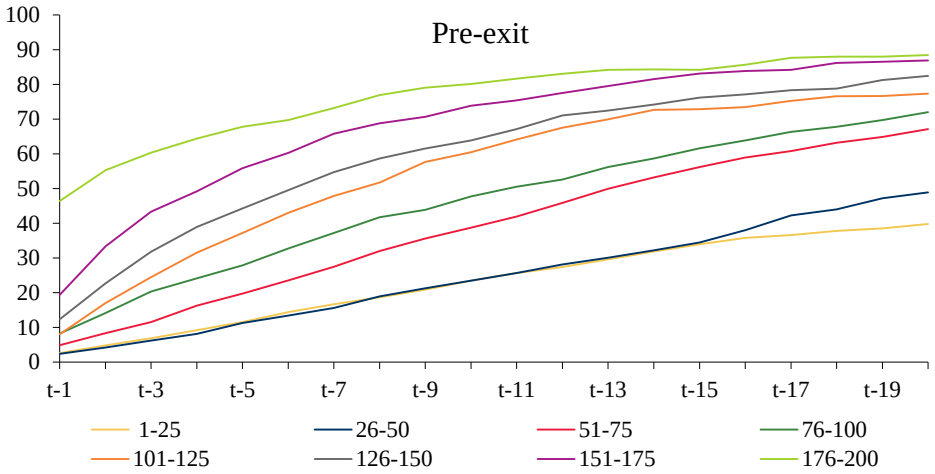


FIGURE 11: Probability of belonging to each class of positions in year $t - x$, given that in year t the firm left the ranking, with x varying between 1 and 20 years

Notes: For each of the 20 transition time horizons considered, the values correspond to those of the column highlighted in red (see Tables A.1, A.2 and A.3 as examples). In each year, the firms that belong to the "Out" category correspond to those that have at least once been included in the ranking of the 200 largest firms, but do not belong to the ranking in that year. In each year, 2.6% of the firms that were in the top 25 in the previous year leave the rank (Table A.1). This probability increases to 23.4% for a 10-year transition horizon (Table A.2) and to 39.8% for a 20-year transition horizon (Table A.3).

The previous results show that the probability of firms changing classes in the ranking is higher for those placed in the lower end and changes are mainly downwards as time goes by. A complementary analysis relies on the net changes in each class (Figure 12). For each class of ranking positions, the net changes correspond to entries minus exits, disaggregated by direction of the move – from upper classes, lower classes or outside the ranking. By construction, since the number of firms in each class is fixed, the net moves cancel out (their sum is zero in all classes). However, it is relevant to notice that the number of upward and downward movements in the ranking is not necessarily symmetric, depending instead on the magnitude of the movement. For example, an upward movement of four classes by a single firm pushes four other firms to the class immediately below. This effect explains why rises are less probable than falls or exits in our database and it is also present when firms enter the ranking (sometimes to intermediate positions). Figure 12 shows that direct entries to intermediate classes are relevant (positive red bars), classes are fed by net moves from upper ones (positive blue bars) and they feed the lower ones (negative yellow bars). Classes [26-50] and [101-125] are those where the contribution to net entry is mostly associated with firms moving from outside the ranking. In addition, in the lowest class [176-200] there is a net move from upper classes that adds to a large number of firms that exit the ranking, movements which are not compensated by entries coming from outside the ranking.

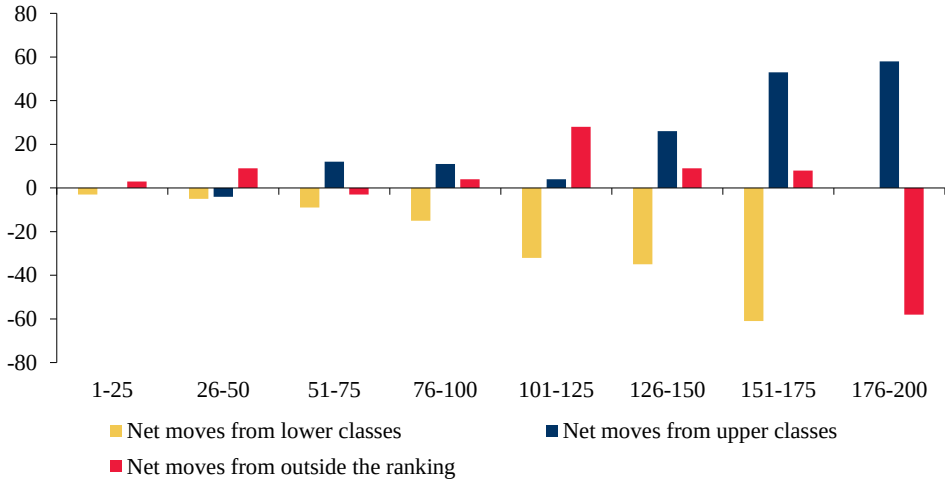


FIGURE 12: Net changes between classes of the ranking from t to $t+1$

Notes: Net changes represent the difference between the number of firms that enter a class of positions (from other classes or from the "Out" category) and the number of firms that exit the same class of positions (to other classes within the ranking or by exiting the ranking). Therefore, "Net changes from lower/upper classes" correspond to entrances from lower/upper classes minus exits to lower/upper classes and "Net changes from outside the ranking" correspond to new entries in the top 200 minus exits from the top 200.

3.2. Nonparametric analysis of duration

In this subsection we use duration analysis methods to estimate firms' probability of remaining in the ranking of the 200 largest firms ("survive") after different time intervals. The event of interest corresponds to firm's exit from the ranking (failure event). In addition, we compare the "survival" experiences across different sectors of activity and size classes.

3.2.1. Methodology and sample characterization

Considering T a non-negative variable denoting the time elapsed between firm entry and exit of the ranking, the survival function is thus represented by:

$$S(t) = 1 - F(t) = Prob(T > t) \tag{1}$$

where the F(t) is the cumulative distribution function for T. The survival function reports the probability of a firm remaining in the ranking beyond t, i.e., the probability that there is no exit prior to t.[4] The most common nonparametric estimate of the survival function is the Kaplan-Meier estimator (López-Garcia and Puente 2006).

For a dataset with k distinct failure times observed in the data, t_1, ...,t_k, the Kaplan and Meier (1958) estimate at any time t is given by:

4. The survival function is a monotone non-increasing function of time. The function is equal to 1 at $t = 0$ and decreases towards 0 as t goes to infinity (Cleves *et al.* 2010).

$$\hat{S(t)} = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right) \tag{2}$$

where n_j is the number of firms at risk (those that remain in the ranking) at time t_j and d_j is the number of failures (those firms that left the ranking) at time t_j. The product is taken over all observed failure times, departing from time t.

Since this estimator is a step function, the estimate of the pth percentile of survival horizons, t_p, is given by:

$$\hat{t_p} = min\left\{ t_i | \hat{S(t_i)} \leq 1 - \frac{p}{100} \right\} \tag{3}$$

for any p between 0 and 100, as described by Cleves *et al.* (2010).

For this estimation, several procedures had to be implemented over the original database. Firstly, firms that were already in the ranking in the first year observed (1981) were excluded, i.e., only firms that entered the ranking in 1982 or later were considered for this specific analysis. Firms that were in the database in 1981, and thus discarded from the sample, represent 24.0% of the total number in the database (835 firms). Out of these 200 firms, 7% belong to the ranking over the entire time horizon (1981-2018).

Secondly, firms with two or more one-year gaps, i.e., those that leave the ranking at least two times and re-enter, and firms with a gap greater than one year were dropped. Firms with two or more one-year gaps represent 4.9% (41 firms) of the total and those with a gap greater than one year correspond to 8.3% (69 out of 835 firms). For firms absent from the ranking only during one year, say year t, we assume that they remain in the ranking and attribute for that year the average of the ranking positions in $t-1$ and $t+1$. These gaps represent only 0.7% of the observations in our database (7600 observations) and are associated to 54 firms.

Thirdly, we assume that a firm does not survive in year t if it is absent from the sample in year $t+1$. Since the last year of the sample is used to identify the firms that exit the ranking in 2017, we restricted the sample to those that entered the ranking between 1982 and 2017 (only 5 firms entered the ranking in 2018). Overall, the sample used in this section takes information for 520 firms over the years from 1982 to 2018 (3583 observations).

3.2.2. Survival functions

The results of the Kaplan-Meier survival estimate for the firms in the sample are plotted in Figure 13. The maximum duration in the ranking of the 200 largest firms is 36 years. Approximately 74.6% of the firms remain in the ranking one year after entry and the estimated median duration is 4 years, meaning that 50% of the firms are expected to remain in the ranking for 4 or less years. After 36 years in the ranking, only about 15.7% of firms "survive".

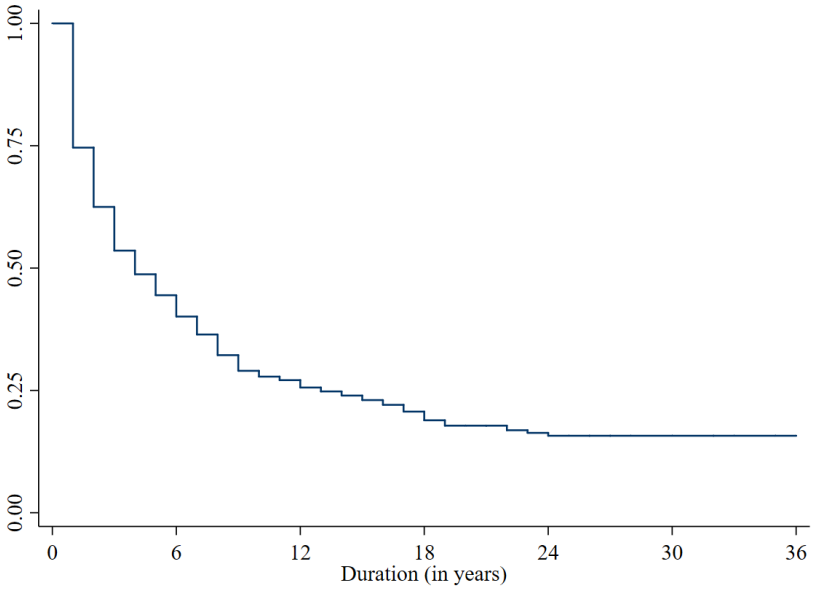


FIGURE 13: Kaplan-Meier survival function for the total sample

The "survival" in the ranking differs according to the sector in which the firm operates (Figure 14).[5] Firms operating in the sector "Electricity and water" have the highest survival probabilities up to the 25th year.[6] The median duration is higher in firms operating in the sectors "Trade, accommodation and food services" (7 years), "Other activities" (5 years) and "Transportation and communication"(4 years). By contrast, firms operating in the sectors "Construction" and "Industry" are those with lower median duration (2 and 3 years, respectively). Nevertheless, after 36 years, only about 21.2% of firms in the "Trade, accommodation and food services" sector remain in the ranking.

In addition, as expected, the smallest firms (in the 1st quartile of the distribution) are less likely to remain in the ranking (Figure 15).[7] By contrast, the largest firms (in the 4th quartile of the distribution) clearly have the highest survival probabilities after 2 years and up to the 35th year in the ranking. The estimated median duration for the smallest firms is 3 years, in sharp contrast with 7 and 8 years for the intermediate classes, respectively.[8] Nevertheless, after 36 years, only about 25.5% of firms in the 2nd quartile of the distribution remain in the ranking.

This additional set of results confirms the conclusions of the sections above regarding the resilience of specific firms in the ranking. The largest firms, which are by construction

5. Both Log-rank and Wilcoxon tests allow for the rejection of the hypothesis of equal survival among sectors.

6. For the sector "Electricity and water", it is not possible to estimate the median duration because the survival function becomes flat in $S(t) = 0.65$, i.e., 65% of the firms in this sector have not "failed" yet.

7. As for sectors, the tests allow for the rejection of the hypothesis of survival equality among classes of firms' size.

8. For the largest firms, it is not possible to estimate the median duration because the survival function has become flat at $S(t) = 0.51$, i.e., more than 50% of the largest firms have not "failed" yet.

those in the top classes of the ranking, have a much higher likelihood of remaining in top positions, i.e., current success seems to enhance future success.

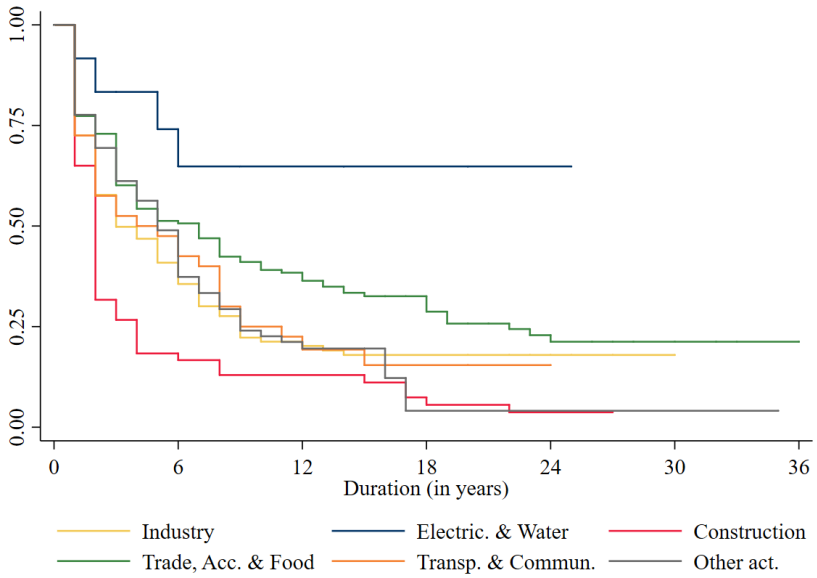


FIGURE 14: Kaplan-Meier survival function by sector

Notes: The sectors presented in this figure correspond to aggregations of NACE Rev.2 sections: "Industry" (sections B and C), "Electricity and water" (sections D and E), "Construction" (section F), "Trade, accommodation and food services" (sections G and I), "Transportation and communication" (sections H and J).

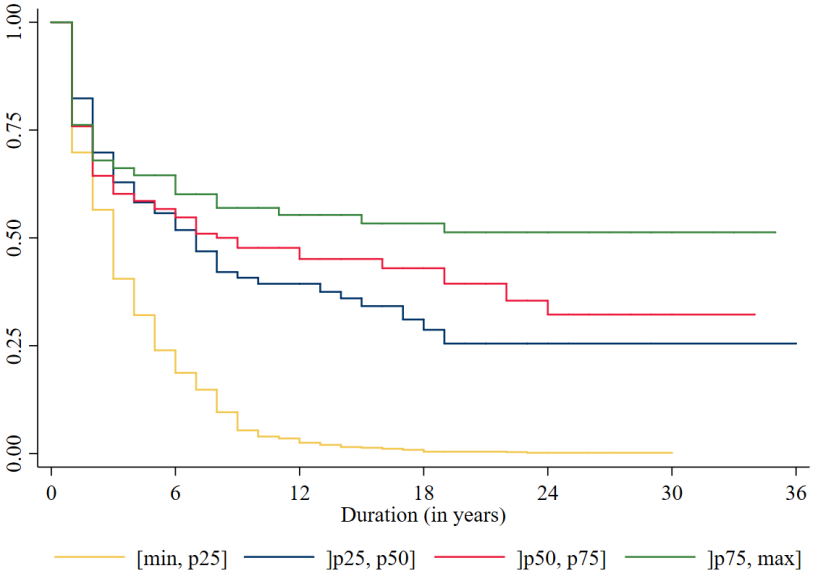


FIGURE 15: Kaplan-Meier survival function by firms' size

Note: Each class of firms' size was defined according to the distribution of turnover in each year. For instance, a firm belongs to the upper class in a year if its turnover was higher than the 75th percentile of the turnover distribution of all the firms in the database in that year. For this analysis, it was considered the modal size class for each firm.

4. Final remarks

This article uses a new database that identifies the top Portuguese firms in the last four decades according to their annual turnover, and establishes some stylized facts regarding their dynamics and survival in the ranking. The results are obtained by computing transition matrices between classes of positions in the ranking for different time horizons and by estimating survival functions.

The empirical literature on firms' demography has provided a rich set of results. However, there is limited evidence relatively to the dynamics of top firms in very long periods of time. We conclude that there is more stability for firms in the top positions of the ranking, showing that size is associated to resilience. In addition, the probability of rising in the ranking in different time horizons for firms in all classes is lower than the probability of falling or exiting. The fact that, on average, the rise in the ranking is harder than the fall is not contradictory with stability in top positions. Although all firms face a sizable risk of dropping out of the ranking, those that have reached the highest positions are comparatively more stable than other top firms that lay in secondary positions. These results are corroborated by survival estimates.

The obstacles to firm growth and their resilience in top positions are important aspects from the perspective of public policies. The rise of firms can be made difficult by different types of regulatory burdens or restrictive competition practices. The fall of firms can be the result of inadequate business models or triggered by unexpected events like technological transformations that turn existing products outdated or by the transfer of a firm's control between generations when management is not separated from property. The analysis and quantification of the determinants of the rise and fall of top firms is a promising avenue for further research.

References

Altomonte, Carlo, Giorgio Barba Navaretti, Filippo Di Mauro, and Gianmarco Ottaviano (2011). "Assessing competitiveness: how firm-level data can help." Tech. rep., Bruegel Policy Contribution.

Audretsch, David B, Enrico Santarelli, and Marco Vivarelli (1999). "Start-up size and industrial dynamics: some evidence from Italian manufacturing." *International Journal of Industrial Organization*, 17(7), 965–983.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen (2017). "The Fall of the Labor Share and the Rise of Superstar Firms." CEP Discussion Papers dp1482, Centre for Economic Performance, LSE.

Cabral, Sónia, Carlos Melo Gouveia, and Cristina Manteu (2020). "The granularity of Portuguese firm-level exports." *Banco de Portugal Economic Studies*.

Cleves, Mario, William Gould, Roberto Gutierrez, and Yulia Marchenko (2010). *An Introduction to Survival Analysis using Stata*. Stata press.

Dunne, Timothy, Mark J Roberts, and Larry Samuelson (1989). "The growth and failure of U. S. manufacturing plants." *The Quarterly Journal of Economics*, 104(4), 671–698.

Farinha, Luísa, Sónia Félix, and João Santos (2019). "Bank funding and the survival of start-ups." *Banco de Portugal working papers*.

Gabaix, Xavier (2011). "The granular origins of aggregate fluctuations." *Econometrica*, 79(3), 733–772.

Kaplan, Edward and Paul Meier (1958). "Nonparametric estimation from incomplete observations." *Journal of the American Statistical Association*, 53(282), 457–481.

López-Garcia, Paloma and Sergio Puente (2006). "Business demography in Spain: determinants of firm survival." *Banco de Espana Research Paper No. WP-0608*.

Mata, Jose and Pedro Portugal (1994). "Life duration of new firms." *The journal of industrial economics*, pp. 227–245.

Mata, José and Pedro Portugal (2002). "The survival of new domestic and foreign-owned firms." *Strategic Management Journal*, 23(4), 323–343.

Mata, José, Pedro Portugal, and Paulo Guimaraes (1995). "The survival of new plants: Start-up conditions and post-entry evolution." *International Journal of Industrial Organization*, 13(4), 459–481.

OECD (2018). "Market Concentration: Issues paper by the Secretariat." Issues paper, Organisation for Economic Co-operation and Development.

Schumpeter, Joseph (1911). *Theori der Wirtschaftlichen Entwicklung*. Berlin: Duncker and Humbolt.

Schumpeter, Joseph (1942). *Capitalism Socialism and Democracy*. New York: Harper.

Wagner, Joachim (1994). "The post-entry performance of new small firms in German manufacturing industries." *The Journal of Industrial Economics*, 42(2), 141–154.

Appendix A

		t+1								
		1-25	26-50	51-75	76-100	101-125	126-150	151-175	176-200	Out
	1-25	87.0	9.3	0.5	0.2	0.1	0.0	0.2	0.0	2.6
	26-50	9.7	72.8	13.7	1.0	0.1	0.3	0.0	0.0	2.4
	51-75	0.1	12.9	60.9	17.2	2.8	0.9	0.3	0.1	4.9
	76-100	0.1	0.8	16.3	52.4	16.0	4.4	1.1	0.6	8.2
t	101-125	0.0	0.4	2.6	15.6	44.5	21.4	5.5	1.9	8.0
	126-150	0.0	0.5	0.8	3.2	19.7	36.0	21.6	5.8	12.3
	151-175	0.1	0.0	0.4	1.1	3.7	17.7	34.3	23.4	19.4
	176-200	0.0	0.0	0.2	0.6	2.1	5.9	16.8	28.0	46.4
	Out	0.1	0.1	0.2	0.3	0.4	0.5	0.8	1.6	95.9

TABLE A.1. Transition matrix with 1 year time horizon

Notes: The rows reflect the initial category (class of positions in the ranking), and the columns reflect the category after 1 year. For instance, each year, 87% of the top 25 firms remained in the top 25 in the next year. The remaining 13% moved to a lower position (10.4%) or left the ranking (2.6%).

		t+10								
		1-25	26-50	51-75	76-100	101-125	126-150	151-175	176-200	Out
	1-25	48.4	14.7	7.4	2.4	1.9	0.3	1.3	0.1	23.4
	26-50	13.3	26.9	15.1	11.3	5.4	2.3	1.6	0.7	23.4
	51-75	5.1	13.6	10.6	12.0	6.9	5.0	5.1	3.0	38.7
	76-100	2.1	5.6	9.0	12.0	9.7	6.4	3.7	3.7	47.7
t	101-125	1.4	2.6	4.1	7.1	5.4	9.1	5.4	4.3	60.4
	126-150	0.6	1.7	4.0	4.9	5.9	6.9	6.7	5.6	63.9
	151-175	0.7	0.3	1.9	2.3	5.9	5.0	5.0	5.1	73.9
	176-200	0.3	0.6	1.4	2.1	3.7	2.9	4.6	4.3	80.1
	Out	1.1	1.3	1.8	1.8	2.2	2.4	2.6	2.9	83.8

TABLE A.2. Transition matrix with 10 years time horizon

Notes: The rows reflect the initial category (class of positions in the ranking), and the columns reflect the category after 10 years. For instance, each year, 48.4% of the top 25 firms remained in the top 25 after 10 years. The remaining 51.6% moved to a lower position (28.1%) or left the ranking (23.4%).

		t+20								
		1-25	26-50	51-75	76-100	101-125	126-150	151-175	176-200	Out
	1-25	29.1	10.7	7.6	5.1	2.0	3.1	1.8	0.9	39.8
	26-50	5.1	17.1	7.1	6.4	5.1	4.7	3.1	2.4	48.9
	51-75	4.2	4.7	4.4	7.1	5.1	3.6	1.8	2.0	67.1
	76-100	1.1	6.7	7.3	2.9	1.8	2.2	2.9	3.1	72.0
t	101-125	1.8	3.1	2.2	3.1	3.1	2.0	3.6	3.8	77.3
	126-150	1.3	1.3	1.3	3.1	3.1	2.9	2.9	1.6	82.4
	151-175	0.9	0.7	0.9	2.0	3.1	2.4	1.1	2.0	86.9
	176-200	0.4	0.2	0.2	1.8	2.2	2.7	1.8	2.2	88.4
	Out	2.2	2.2	2.7	2.7	2.9	3.0	3.2	3.2	77.8

TABLE A.3. Transition matrix with 20 years time horizon

Notes: The rows reflect the initial category (class of positions in the ranking), and the columns reflect the category after 20 years. For instance, each year, some 29% of the top 25 firms remained in the top 25 after 20 years. The remaining 71% moved to a lower position (31.2%) or left the ranking (39.8%).

Appendix B

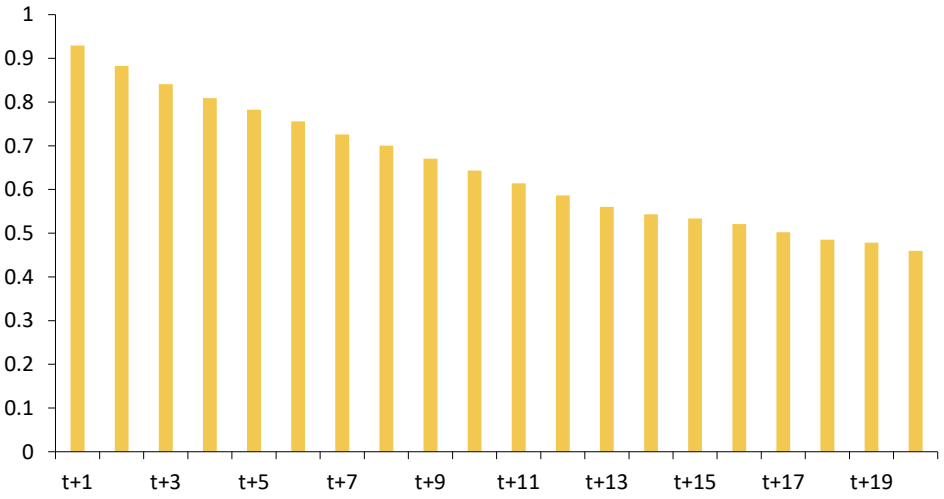


FIGURE B.1: Spearman's rank correlation coefficient between the firms' class in year t and year $t + x$