BETA
Bureau
d'économie
théorique
et appliquée

Documents
de travail

«Entry, Exit and Productivity: Evidence from French Manufacturing Firms »

Auteur

Enrico De Monte

Document de Travail n° 2020 – 07

Février 2020

Bureau d'Économie Théorique et Appliquée
BETA

www.beta-umr7522.fr

@beta_economics

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Entry, Exit and Productivity: Evidence from French Manufacturing Firms*†

Enrico De Monte‡

February 28, 2020

Abstract

This paper analyzes productivity dynamics based on French firm-level data covering nine key 2-digit industries for the period 1994 - 2016. I estimate firm-level productivity through the estimation of a translog production function and investigate the following main aspects: (i) aggregate productivity change with firm entry and exit by applying the Dynamic Olley-Pakes Productivity Decomposition (DOPD), (ii) firms' ability to improve productivity and productivity persistence, and (iii) productivity differences between different firm groups such as survivors, entrants and exitors as well as small, medium and big firms by applying the concept of stochastical dominance. My results show that aggregate productivity has increased for most the considered 2-digit industries and that in many cases surviving firms' have contributed significantly to these positive improvements. Entering firms contribute in many cases positively to aggregate productivity while the contribution of exitors shows varying signs. Furthermore, I find that firms' reveal a high degree of productivity persistence. Analysing productivity difference between firm groups the results suggest that the productivity distribution of surviving firms stochastically dominates the distribution of entering and exiting firms. Surprisingly, the results reveal that big firms do not stochastically dominate the productivity distribution of small firms.

Keywords: production function estimation, productivity decomposition, technological change, productivity differences, firm entry and exit.

JEL Classification: C13; C14; D24; D30; L60; O47

*I would like to thank Bertrand Koebel, Rabah Amir, Phu Nguyen-Van, Anne-Laure Levet, Gabriele Rovigatti and the participants of seminars and conferences in Champ-sur-Marne (FCBA), Nancy (Seminar, BETA Nancy) and Lyon (ADRES, Doctoral Conference) for their helpful comments. I would also like to thank FCBA and ONF for their precious support during the realization of the project.

†This project is funded by FCBA (French Institute of Technology for Forest-based and Furniture Sectors), ONF (National Forests Office) and Chair Gutenberg (research fund of the eastern region of France).

‡E-mail: e.demonte@unistra.fr. Address: Université de Strasbourg, BETA, 61 Avenue de la Forêt-Noire, 67085 Strasbourg, France.

1 Introduction

The economic crisis in 2007 has impacted the French economy persistently. The data show that for the most important French manufacturing industries the production level has significantly dropped during the economic and financial distress and has not yet attained its production level before the crisis. The decrease in the French manufacturing industries' production level is also reflected in a considerable decrease in the number of firms, implying higher exit than entry rates.

I am interested in investigating how firms' entry-exit dynamics are related to individual firms' productivity as well as to the industries' aggregate productivity trajectories. For this purpose I analyze firm-level data of the universe of firms active in the French manufacturing industry between 1994 and 2016 and estimate firm-level productivity based on a translog production function.

More precisely, to shed light on the relation between firms' productivity and their status of either survivor, entrant or exitor I study productivity with respect to the following aspects: (i) the evolution of aggregate productivity related to firm entry and exit, (ii) productivity persistence and (iii) productivity differences between different firm groups such as survivors, entrants and exitors. I am also interested in investigating productivity differences between firms belonging to different size groups such as small, medium and big firms.

As extensively discussed in the literature, firm-level productivity itself can be driven by many factors, such as managerial practice, higher-quality inputs factors, R&D activity, firm structure and product innovation (Bartelsman and Doms, 2000; Syverson, 2011). The analysis of drivers of firm productivity is beyond the scope of that work. Rather than analyzing determinants of single firms' productivity level this paper investigates more general patterns such as aggregate productivity growth, and productivity differences among various firm groups. Firm selection, i.e. the selection of firms that survive in the market and the resulting number of firms, is a widely discussed issue in the field of industrial organization. For instance, the model presented in Jovanovic (1982) explains market selections by efficient firms that grow and survive and inefficient ones that shrink and exit, where the only source of heterogeneity is generated from efficiency differences in firms' variable costs. In the industry model presented by Hopenhayn (1992) firms are exposed to idiosyncratic productivity shocks, that follow a first order Markov process. That is, entering firms' initial productivity levels are drawn from the same distribution and active firms with a higher productivity in the current period are more likely to be affected by a positive productivity shock in the future period. Firms whose productivity level falls below a certain lower threshold exit the market. I aim to empirically investigate whether there are significant productivity differences between those firms that survive, enter and exit the market. Also, as Hapenhayn's model describes, initial productivity differences may be important for firms future performance.[1] Similarly in other models of industrial dynamics, such as Ericson and Pakes (1995), productivity plays a key role for firm survival and, hence, the provision of descriptive statistics and the derivation of stylized facts can be viewed as a first step for a better understanding of the process of firm selection.

The first aspect to be investigated in this study, the evolution of aggregate productivity and the contribution of firm entry and exit, aims to provide information on the trajectory of aggregated productivity and on whether the manufacturing industries moves to higher allocative efficiency. Generally, aggregate productivity is measured as a weighted average composed of firms' market shares (weights) and the corresponding productivity levels. If firm exit occurs, the market shares of failing firms are recovered by either entering or surviving firms. Likewise, market shares may wander from less productive continuing firms to more productive continuing or entering firms. In the empirical literature this is called allocative efficiency (of production), extensively discussed by Baily et al. (1992), Griliches and Regev (1995), Pavcnik (2002) and Polanec (2004). In this study I rely on the "Dynamic Olley Pakes Decomposition" (DOPD) developed by Melitz and Polanec (2015). Here aggregated productivity growth is decomposed into the contribution of surviving firms as well as the contribution of entering and exiting firms. Moreover, aggregate productivity of surviving firms is further decomposed into the part of productivity growth induced by firms' individual improvement in productivity (within productivity change) and into that part associated with market reallocations (between firm productivity change). As pointed out by Haltiwanger (2011) the measure of allocative efficiency is helpful to assess whether an economy is "well-functioning". According to Haltiwanger, an economy is allocative efficient from a static viewpoint if more productive firms produce more. It is allocative efficient from a dynamic perspective, if production is reallocated from less to more productive firms.

The second objective of the paper is the analysis of persistency in firms' productivity ranking. Empirically, many studies have revealed strong dispersion of productivity within a given industry

[1] Also see Esponda and Pouzo (2019) for a recent extension of Hapenhayn's model to the case of heterogeneous firms.

(Syverson, 2004). Also, empirical work has shown that firms tend to reveal a high degree of persistence with respect to their productivity ranking (Haltiwanger, 2011; Foster et al., 2006, 2008). I aim to provide information on this discussion for the French case. The third objective of the paper is to assess productivity differences between entering, surviving and exiting firms. Many models describing industrial dynamics, i.e. the ongoing process of firm entry and exit, are able to generate observed patterns in firm dynamics supposing that firm exit is driven by a lower productivity level compared to other competitors. I examine whether this assumption is valid for the case of firms active in French manufacturing industries. For this purpose I follow the empirical strategy in Fariñas and Ruano (2005) who analyze Spain manufacturing industries. That is, I apply the concept of stochastical dominance by comparing the productivity distributions of the considered firm groups. The investigation of the relation between firm productivity and firm selection is crucial to understand whether the economy moves towards higher efficiency. This view is related to the Schumpeter theory referring to creative destruction, where more efficient firms are supposed to replace less efficient firms. It is therefore interesting to investigate whether such dynamics can also be observed in the case of the French manufacturing industries, especially in the light of the economic distress during the past decade.

My results show that for most industries aggregate productivity is mainly driven by productivity improvements of surviving firms (within change). Also, I find that the contribution of aggregate productivity changes due to shifts in the market shares of surviving firms (between change) is mainly negative, indicating inefficient dynamic allocation (Haltiwanger, 2011). Furthermore, for many industries and years, firm entry is positively related to aggregate productivity growth. Firm exit, instead, shows varying signs with respects to its contribution to aggregate productivity. However, note that I find substantial differences among industries with respect to aggregate productivity dynamics and the contributions of the three firm groups, survivors, entrants and exitors. My study also shows that firms' have a high degree of persistency in productivity, meaning that firms' current productivity level strongly determines their future productivity level. The investigation of productivity differences reveals the following results: survivors have higher productivity levels compared to entering and exiting firms; entrants have higher productivity compared to exitors; and surviving entrants have higher productivity compared to failing entrants;[2]

The paper is organized as follows. Section 2 reviews more in detail the related literature; Section 3 presents the data and descriptive statistics; Section 4 introduces the analytical framework; Section 5 describes the empirical results; Section 6 discusses limits of my study; and Section 7 concludes.

2 Related literature

Productivity dynamics is a very well documented area in the field of empirical industrial organization. The main objectives of this work is the study of aggregate productivity for the case of the French manufacturing industries as well as to investigate productivity differences between entering, surviving and exiting firms as well as differences among size groups. In the following I will present the related literature with respect to these aspects.

2.1 Aggregate productivity

Aggregated productivity can be viewed as a weighted average of individual firms' productivity level, given by

$$\Omega_{It} = \sum_{n \in I}^{N_{It}} s_{nt}\omega_{nt}, \qquad n = 1, \cdots, N_{It}, \tag{1}$$

where Ω_{It} denotes the aggregate productivity level of a given industry I at point t, s_{nt} is the markets share of a firm n and ω_{nt} its corresponding individual productivity level. N_{It} denotes industry I's the total number of active firms in t. $\Delta\Omega_{It} = \Omega_{It} - \Omega_{It-1}$, the change in aggregate productivity, varies due to two reasons: (i) through a change of single firm's productivity ω_{nt}, in this case the literature refers to "within-firm" productivity change; (ii) market-share reallocations, i.e. a change in a firm's market share s_{nt}; in this case the literature refers to "between-firm" productivity change. The number of firms in a given industry changes with firm entry and exit over time. Entering and incumbent firms recover market shares from other incumbents and/or

[2]Surviving entrants are firms that are active for at least two consecutive years after entry. Instead, failing entrants are firms that exit the market in the year following the year of entry. For more details, see Section 5.4.4.

exiting firms. In this context it is interesting to measure to which extent entering and exiting firms contribute to the aggregate productivity. Some important studies considering the measurement of these aspects are reviewed in this section.

An extensive study on patterns of productivity dynamics was conducted by Baily et al. (1992) using U.S. plant-level data ranging from 1963 to 1987. They developed probably the first productivity decomposition to measure explicitly the contributions of entrants, survivors and exitors to aggregate productivity growth. Baily et al. (1992) find that entry and exit has a very small contribution to aggregate productivity changes. Instead, within-firm productivity changes represent the largest effect on aggregate productivity growths, whereas reallocation of output shares, i.e. between-firm changes in productivity, has a much smaller but positive impact. The positive contribution of between-firm productivity growth implies that firms doing well in productivity relative to their competitors gain market shares. Analyzing the Israeli manufacturing industry between 1979 and 1988, Griliches and Regev (1995) also find that output reallocation between firms played a much less important role compared to within contribution, i.e. firms' proper productivity growth, which has increased the aggregated productivity at most. Investigating the U.S. telecommunication industry Olley and Pakes (1996) find evidence of an important role of reallocation for aggregate productivity growth. Melitz and Polanec (2015), to my knowledge the most recent approach of productivity decomposition, apply their method on the Slovenian manufacturing industry, covering the period from 1995 - 2000. The study reveals that within-firm growth in productivity carries the largest part of the aggregate productivity growth.[3] Levinsohn and Petrin (1999) investigate Chilean data of firms belonging to the manufacturing industries. They refer to the "real productivity case" if single firms' productivity changes and to the "rationalization case", when firms' market shares are transferred to more productive firms by the process of market selection. They find that the rationalization case is of higher empirical importance when industries gain in productivity. Instead, when industries' aggregate productivity decreases, the real productivity case turns out to be more important. Foster et al. (2001) review many of the methods applied for the decomposition of productivity growth. Comparing the studies they highlight that the within-firm component varies with the business cycle: in phases with strong productivity growth, the within-firm component contributes significantly, whereas in phases with lower growth rates the within contribution decreases. Foster et al. (2006) analyze productivity dynamics in the U.S retail trade sector, covering a period over 10 years. They find that reallocation accounts considerably to productivity growth, i.e. aggregated productivity mainly increased due to the transfer of market shares from low to higher productive firms. They also show that firms are only little mobile in terms of productivity ranking, meaning low/high productive firms tend to stay low/high productive firms, given they have survived.[4] Moreover, their study reveals that less productive firms are much more likely to exit the market compared to high productive firms. Foster et al. (2008) confirm this result, using U.S manufacturing firm-level data, from 1977 to 1997.

In the international trade literature, the concept of productivity decomposition is used to measure the contribution of exporting and nonexporting firms to aggregate productivity growth in order to derive the effect on trade liberalization of a given industry's performance. For instance, Pavcnik (2002) analyzes Chilean firm-level data from manufacturing industries and finds that about 1/3 of the increase in aggregated productivity is associated with individual firms' productivity improvement (within-firm increase in productivity) and about 2/3 to market share shifts from less to more productive firms (between-firm increase in productivity). Moreover, aggregate productivity grew most in sectors in which international trade was prevailing. Similarly, Bernard and Jensen (1999) analyse U.S. data of manufacturing industries and find that roughly 20% aggregate productivity growth is accounted by reallocation of market shares from of less productive non exporting firms towards their more productive exporting competitors.

Generally, the presented studies show that there is no common law which of the productivity components - i.e. within/between productivity improvement of surviving firms as well as the contribution of entering and exiting firms - dominates aggregate productivity growth. Instead, the studies suggest that this depends on the inherent characteristics of a given industry as well as exogenous macroeconomic influences. To give an example, the already mentioned study by Foster et al. (2006) showed that the retail sectors are much more dynamic with respect to entry and exit compared to manufacturing industries. The between firm change in aggregate productivity is much more important than the within change. Foster et al. (2006) explain this latter finding by the fact that in the retail sector output markets are much more flexible, i.e. consumers can easily change suppliers, implying that market shares also change more easily from less to high productivity firms.

[3]See details in Section 4.3.

[4]Also see Baily et al. (1992), who find similar results.

General differences in terms of entry and exit dynamics between the manufacturing and other industries such as the the retail sector may be due to lower entry barriers: the manufacturing industries are typically characterized by higher fixed costs to carry by firms. If fixed costs are sunk they can be linked to entry barriers which in turn explain low entry rates, see Geroski (1995) and the cited literature therein. When considering the within-firm productivity improvements the question of what determines firm productivity is indispensable. There is a large list of factors directly affecting firm-level productivity, to name only a few: managerial practice, higher-quality inputs factors, R&D activity, firm structure and product innovation. The analysis of drivers for individual productivity is beyond the scope of my work, for a detailed discussion see Bartelsman and Doms (2000) and Syverson (2011).

2.2 Productivity differences between firm groups

In well known industry models such as Jovanovic (1982) and Hopenhayn (1992) firms' productivity plays an important role both with respect to market selection and firm size. For instance, in the model presented in Hopenhayn (1992), firms' at their moment of entry draw their initial productivity level, say ω_{nt}^{0} from the same distribution. In future periods, firms' productivity is exogenously updated, following a first-order Markov process. That is, firms with higher initial productivity levels are more likely to receive positive productivity shocks (updates), whereas low productivity firms are more likely to receive relatively small or negative productivity shocks. The model further assumes that there is a productivity threshold, say $\underline{\omega}$, below which firms exit the market at t (firm selection) if their productivity level falls below the threshold at $t-1$. This process should, generally, sustain dynamic allocative efficiency, already mentioned above. This is because entering and/or surviving firms, endowed with a productivity above the threshold $\underline{\omega}$, take over left market shares from the exiting firms, implying that market shares wander from lower to higher productivity firms. Fariñas and Ruano (2005) tested the hypothesis of Hopenhayn's model for Spain manufacturing industries between 1990 and 1997. Their results confirm the model assumptions, finding that surviving firms exhibit higher productivity levels compared to entrants and exitors. They also showed that entrants have higher productivity compared to exitors, for a given cohort. The importance of initial productivity endowment is also shown by the finding that surviving entrants turn out to have higher productivity levels compared to failing entrants.[5] Wagner (2010) replicates the study presented by Fariñas and Ruano (2005) for the German case (for both West and East Germany). Using labor productivity as a measure for firm productivity Wagner's results confirm Fariñas and Ruano (2005) findings. As Wagner (2010) highlights, the results about productivity differences with respect to the firms groups survivors, entrants and exitors underline the Schumpeter theory of creative (or constructive) destruction, i.e. firms with higher productivity replace firms with lower productivity. As already mentioned, in the industry models presented by Jovanovic (1982), Hopenhayn (1992) or Ericson and Pakes (1995), the only source of heterogeneity is given by firms' level of productivity. For this reason, firm sizes should generally reflect firms' level of productivity, i.e. higher productivity firms should also be larger. A positive correlation between firm productivity and firm size is found by Leung et al. (2008) for Canadian manufacturing and non-manufacturing industries. They find for both sectors a strong relation between firm size and firm productivity (measured both in labor productivity and total factor productivity (TFP)), with a large productivity gap between small and large firms. Van Biesebroeck (2005) found evidence for higher labor productivity of large firms active in African manufacturing industries. Instead, when considering total factor productivity, the difference seem to vanish, i.e., larger firms do not systematically show a higher total factor productivity compared to smaller firms.

3 Data and descriptive statistics

I analyse French firm-level data containing the universe of firms active in different 2-digit manufacturing industries. For this purpose I combine the (fiscal) data bases FICUS and FARE covering the periods 1994-2007 and 2008-2016, respectively. The data bases contain detailed information about firms' reports in balance sheets and income statements. Note that, in 2008 the French institute for statistics and economic studies (INSEE) made significant changes with respect to the industry nomenclature firms belong to. In both data bases, the principal industry identifier is on the 4-digit level, where in FICUS industries industries were differently labelled compared to FARE.

[5] Note that Fariñas and Ruano (2005) analysing data from 1990 to 1997 define surviving and failing entrants for five cohorts, 1990 - 1994. For each cohort surviving entrants are identified by those firms surviving until 1997, whereas failing entrants are those firms exiting the market before 1997.

In order to guarantee consistency in the industry nomenclature I manage to use throughout the whole period, 1994-2016, the same industry nomenclature. This is important especially for firms that have exit the market before 2008, and for which it is not known for sure to which industry they would have belonged to in FARE. For a more detailed description of the construction of the data set, see Appendix A.

3.1 Variables

Since my prior interest is to estimate firm-level productivity through the estimation of a production function, I describe in the following the required variables for this purpose. Beginning with firms' gross output I use firms' total production, which is the sum of firms' sales, stocked production and capitalized production. Furthermore, as firms' capital stock I use their amount of tangible assets, labor is measured by the number of employees and intermediary products consumption by the sum of the expenditures for raw and intermediary materials. All variables are deflated by the corresponding 2-digit industry price index. It is noteworthy that in FICUS firms with zero employees (i.e. self-employees) are not explicitly observed, even though firms may temporarily (for economic reasons) report zero employees. In FARE, instead, these firms are observed, which enlarges considerably the data base. To establish consistency between both data sets firms with zero employees are dropped.[6]

3.2 Number of firms, entry and exit

Since I use fiscal data, firms' report on their balance and income statement is mandatory. However, I also observe some non-report, especially for very small firms. Generally, the number of firms varies in the data through non-report, ambiguous firm status (temporal inactivity) and firm entry and firm exit. Unfortunately, it is not definitely possible to distinguish between non-report, temporal inactivity, and firm exit. For this reason I adopt the following approach to identify firm entry and exit.

Let $a_{nt} \in \{0, 1\}$ be a firm state variable, taking the value 0 in case of inactivity, and 1, if the firm is active. A firm is said to be active at t, if it reports nonzero data for one of the following variables: total production, turnover and/or net profits. In all other cases the firm is supposed to be inactive. Further, survival is denoted by $s_{nt} \in \{0, 1\}$ with $s_{nt} = 1$ if $a_{n,t-1} = a_{nt} = a_{n,t+1} = 1$. Entry is denoted by $e^+_{nt} \in \{0, 1\}$ with $e^+_{nt} = 1$ if $a_{n,t-1} = 0$ and $a_{nt} = a_{n,t+1} = 1$. Exit is denoted by $e^-_{nt} \in \{0, 1\}$ with $e^-_{nt} = 1$ if $a_{n,t-1} = a_{nt} = 1$ and $a_{n,t+1} = 0$. In the literature firm entry and exit is often measured by looking one period ahead (see for instance Blanchard et al. (2014)). I.e., $e^+_{nt} = 1$ if $a_{n,t-1} = 0$ and $a_{n,t} = 1$, and similarly with firm exit. However, measuring entry and exit in this way introduces some ambiguity with respect to the identification of entrants and exitors. This can be seen in Table 1. In the very last row, where the firm is only active in t, it could be considered as an entrant and/or exitor in t. Instead, I prefer to use the alternative convention and consider firms exhibiting an activity sequence as described in the last row of Table 1 as unidentified.

Table 1: Firm status example

Variable activity (0/1)				
$a_{n,t-1}$	a_{nt}	$a_{n,t+1}$	Status in t	Binary firm status variables in t
1	1	1	Survivor	$s_{nt} = 1,\ e^+_{nt} = 0,\ e^-_{nt} = 0$
0	1	1	Entrant	$s_{nt} = 0,\ e^+_{nt} = 1,\ e^-_{nt} = 0$
1	1	0	Exitor	$s_{nt} = 0,\ e^+_{nt} = 0,\ e^-_{nt} = 1$
0	1	0	Not identified	$s_{nt} = 0,\ e^+_{nt} = 0,\ e^-_{nt} = 0;$

Using these definitions, the aggregate number of firms is given by:

$$\begin{aligned} N_t = \sum_{n \in I_t} \mathbf{1}_{[a_{nt}=1]} &= \sum_{n \in I_{t-1}} \mathbf{1}_{[s_{nt}=1]} + \sum_{n \in I_t} \mathbf{1}_{[e^+_{nt}=1]} + \sum_{n \in I_t} \mathbf{1}_{[e^-_{nt}=1]} + \sum_{n \in I_t} \mathbf{1}_{[u_{nt}=0]} \\ &= S_t + E^+_t + E^-_t + U_t \\ &= N_{t-1} + E^+_t - E^-_{t-1} + U_t \end{aligned} \tag{2}$$

[6]Note that in the data I frequently observe that very small firms alter between reporting zero and one employee. See Appendix A for more details on changes in the number of observations of firms with zero and one employee.

where I_t denotes the set of firm IDs included in the data at t. The total number of survivors, entrants, exitors and non-identified firms are denoted by S_t, E_t^+, E_t^- and U_t. The notation $\mathbf{1}_{[A]}$ denotes a dummy variable equal to 1 if the condition A in brackets is satisfied and 0 otherwise.[7]

3.3 Descriptive statistics

I now present descriptive statistics of the underlying data set. Table 2 summarizes some statistics with respect to firm size. The figures represent averages over all years. It can be seen that most firms are within the groups up to 9 employees. The first three size groups represent more than 70% of all firms. Generally the number of firms decreases with firm size, where the group of firms with more than 500 employees represent only 0.7% of all firms. Contrarily, small firms only represent a very small percentage in terms of labor demand: all firms with less than 49 employees demand about 25% of all work force, where the group of firms with more than 500 employees demand about 40% of the work force. The table also shows that firm age is increasing in firm size. Regarding the last three columns that provide insights about firms survival, entry and exit rates. The figures show that small firms are much more dynamic in terms of entry and exit compared to larger firms. I measure an entry (exit) rate for firms with only one employee of about 14.8% (15.0); for firms with more than 100 employees only about 3% (3.5%) and less.

Table 2: Summary Statistics Size

Size	# of firms	Share of firms	# of employees	Share of employees	Age	Survival Rate	Entry Rate	Exit Rate
1	8810	16.7	8810	0.6	12.1	55.1	14.8	15.0
2-4	16846	31.8	48161	3.3	12.3	73.0	9.1	8.7
5-9	12054	22.8	79979	5.4	14.1	77.3	6.6	6.9
10-19	6202	11.7	84376	5.7	17.7	81.1	4.9	5.5
20-49	5031	9.5	160544	10.8	22.7	84.0	3.2	3.9
50-99	1690	3.2	119209	8.1	25.6	84.5	2.9	3.8
100-199	1111	2.1	156271	10.6	26.8	85.1	2.8	3.5
200-499	771	1.5	236276	16.0	27.7	86.3	2.5	2.8
>=500	383	0.7	586439	39.6	28.5	85.9	3.0	2.7
Total	52898	100.0	1480065	100.1	19.7	72.9	9.7	7.8

Analogously, Table 3 shows the same statistics with respect to the 2-digit industries. As can be seen that industry 10 (food processing) is by fare the largest industry, including about 57% of all firms representing a share of employees of about 31%. This industry also shows the highest average firm entry rates, given by 8.8%. Share of firms for the remaining industries are given for 4.7% (textiles), 8.9% (wood products), 2.3% (pulp/paper), 3.9% (chemical products), 1.5% (metals), 9% (machines), 2.9% (automobiles) and 9.1% (furniture). In terms of number of employees the industries "chemical products", "machines" and "automobiles" employ most workers (after industry "food"), with shares respectively given by 12%, 12.8% and 17.8%. Across the industries, the average entry rates vary between 4.4 % (pulp/paper) and 8.6% (furniture). Similarly, the average exit rates varies between 5.2% (pulp/paper) and 9.9% (furniture).

Table 3: Summary Statistics: Industries

Industry[a]	# of firms	Share of firms	# of employees	Share of employees	Age	Survival Rate	Entry Rate	Exit Rate
10	30539	57.7	471781	31.9	13.6	72.3	8.8	8.5
13	2465	4.7	70331	4.8	19.9	77.5	5.8	7.2
16	4729	8.9	70667	4.8	17.3	76.2	6.8	7.2
17	1220	2.3	79510	5.4	21.6	81.2	4.4	5.2
20	2042	3.9	178005	12.0	20.8	78.5	5.7	6.2
24	813	1.5	93098	6.3	20.6	79.8	5.7	5.6
28	4753	9.0	189750	12.8	17.5	76.2	7.3	7.5
29	1548	2.9	263230	17.8	18.2	78.3	6.0	6.2
31	4803	9.1	63711	4.3	14.9	71.3	8.6	9.9
Total	52912	100.0	1480083	100.0	18.3	76.8	6.6	7.1

[a] Industry description: 10 (food), 13 (textile), 16 (wood), 17 (pulp/paper), 20 (chemical products), 24 (metals), 28 (machines), 29 (automobiles), 31 (furniture).

[7] Note that the data does not allow to track mergers and acquisitions. This means that if, for instance, two firms have merged, the emergent firm obtains a new firm id, whereas the former ids from the two merged firms disappear. Statistically, for this case and the underlying data, we observe two firm exits and one entry.

Figure 1 illustrates the times series of the number of firms for different size groups with respect to the number of employees: small firms with less than 10 employees, small-medium firms with 11-50, medium firms with 51-150, medium-big firms with 151-500 and big firms with more than 500 employees. The left plot shows the evolution of the number of firms of the groups "small" and "small-medium", where the solid line represents the groups of small firms. It can be seen that in 2007 there is a significant decrease in the number of firms. This relatively low number of firms in 2015 is linked to firms that are exceptionally registered with zero employees and, hence, drop out of the sample (since I only keep firms with at least one employee). Many dropped firms report in 2016 again one employee (or more), which is the reason for the sharp increase in the number of firms for that year.[8] The group of small-medium firms, represented by the dashed line, also exhibits a negative trend from 2005, however, the negative trend is much more modest compared to the decrease in the number of small firms. Considering the figure on the right hand side, all firm groups, "medium" (dotted line), "medium-large" (dashed-dotted line) and "large" (dashed line) show, after a their highest level between 2000 and 2005 a significant ongoing negative trend until 2016.

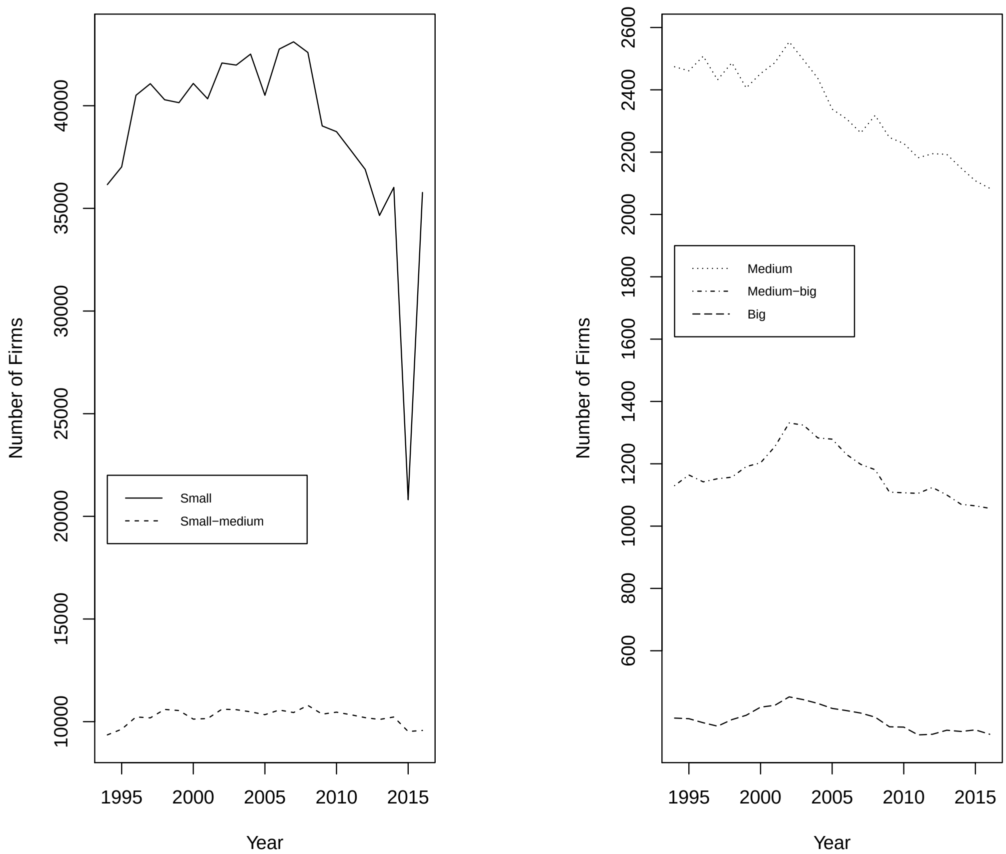

Figure 1: Evolution of the number of firms for different size groups.
Note: Firms size groups with respect to the number of employees: small (1-10), small-medium (11-50), medium (51-150), medium-big (151-500), big (>500).

Figure 2 summarizes throughout the whole period and over all industries the evolution of the number of firms (on the left y-axis) as well as the entry and exit rates (on the right y-axis). Again there is a negative trend in the number of firms after 2007 (represented by the dashed line). This pattern is also reflected by the entry and exit rates, represented by the solid and dotted lines, respectively. While before 2007 these rates fluctuate around each other, from 2007 the exit rate lies continuously over the entry rate, translating into excess exit.

[8] See Appendix A for more details with respect to firms shifting between zero and one employee.

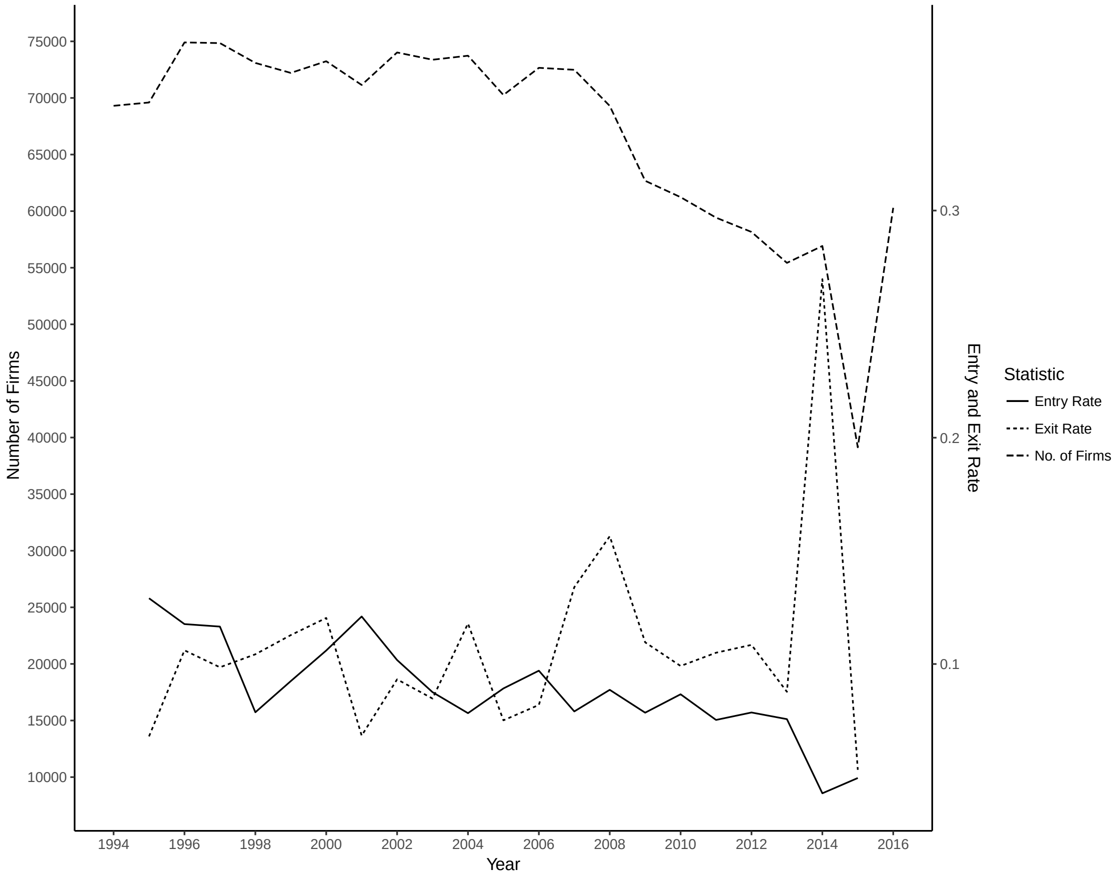

Figure 2: Firm dynamics

To complete the description of the data Figure 3 illustrates the change over time of the aggregate of the input factors and output. For each variable, the time series represents the index value (normalized to 100 in 1994) of the aggregate over all firms. It can be seen that total production, represented by the solid line (closely followed by the dotted line for intermediary products), experiences a sharp decline during the time of the financial and economic crisis between 2007 and 2009, and seems to slightly recover until 2016, ending up with a production level of 123.1% compared to 1994. Aggregate capital, instead, rose constantly and reached in 2016 a level 85.7% above the initial one. Contrarily, aggregate labor has decreased from 2002 on, with a level of only 90.5% compared to 1994.

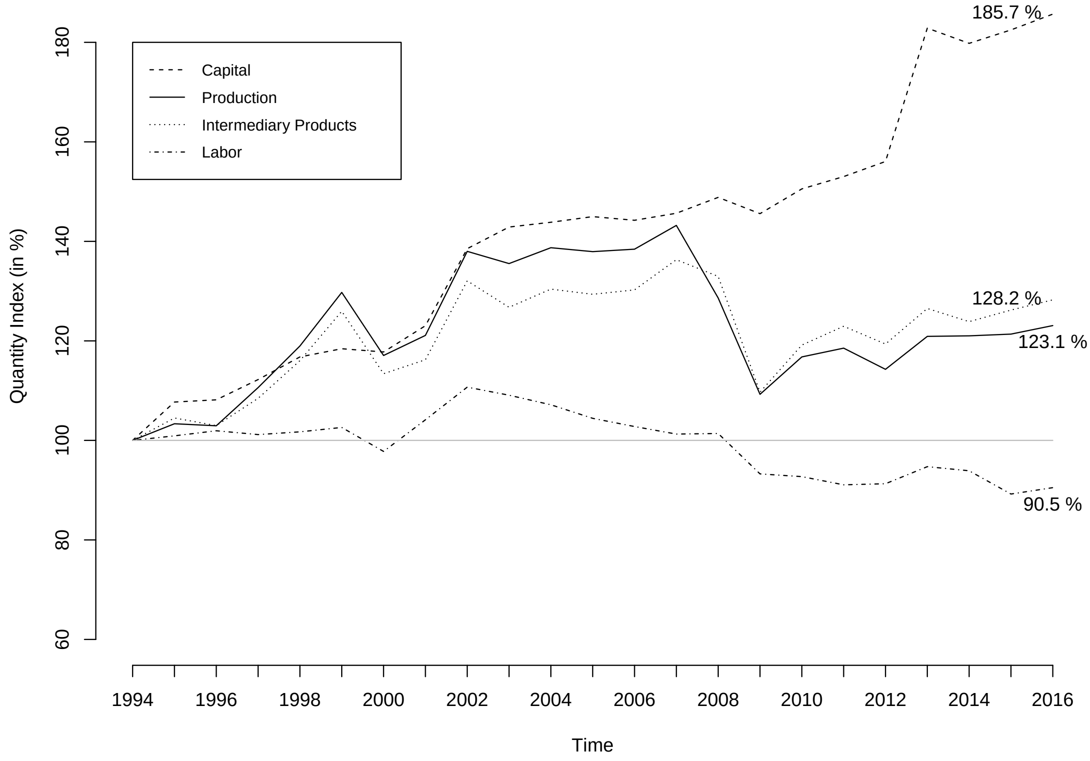

Figure 3: Evolution input quantities and production

4 Analytical framework

The objectives of the study is the investigation of firm-level productivity with respect to three aspects: (i) studying aggregate productivity change by decomposing total factor productivity (TFP) into the contribution of surviving, entering and exiting firms, (ii) identifying the degree of persistence of productivity and (iii) comparing the productivity differences between entrants, survivors and exitors. For this purpose, the first step is to obtain consistent estimates of the production function, from which firm-level productivity is recovered. In the following I describe the adopted methodology to estimate the production function as well as the analytical framework for the mentioned points (i) - (iii).

4.1 Productivity estimation

A firm n's productivity measures how efficiently the firm transforms production inputs into output compared to its competitors. Formally, productivity is integrated in the production technology which is described by a production function. For my purpose I suppose a gross output translog production technology, given by

$$y_{nt} = \alpha_0 + \sum_i \alpha_i x_{nt}^i + \frac{1}{2} \sum_{ij} \alpha_{ij} x_{nt}^i x_{nt}^j + \omega_{nt} + \epsilon_{nt}, \tag{3}$$

where y_{nt} denotes the logarithm of a firm's gross output production and x_{nt}^i with $i = (k, l, m)$ denotes the log of input factors capital, labor and intermediary products (materials). ω_{nt} represents the log-level of productivity, known to (or anticipated by) the firm but unknown to the econometrician. ϵ_{nt} represents an iid shock. Note that, as common in the production function literature, I suppose that firms' capital stock evolves according to $K_{nt} = \kappa(K_{nt-1}, I_{nt-1})$, where $K_{nt} = \exp\left(x_{nt}^k\right)$ and I_{nt} denotes a firm's amount of investments. This timing assumption implies that capital is fixed when a firm observes the innovation in productivity ω_{nt}. Instead, labor inputs are supposed to be flexible, and hence adjustable with respect to ω_{nt}. As has been extensively discussed in many studies such as Olley and Pakes (1996) (OP, henceforth), Levinsohn and Petrin (2003) (LP, henceforth), Ackerberg et al. (2015) (ACF, henceforth) and Wooldridge (2009) a crucial difficulty to deal with when estimating production functions consists in the endogeneity of the explanatory variables, arising when a firm chooses its flexible inputs (here x^l_{nt}, x^m_{nt}) as a function of the productivity shocks ω_{nt}. To circumvent the endogeneity problem OP were the first proposing a two stage estimator and using firm investments as proxy variable to control for unobserved heterogeneity. The LP approach suggests to use materials as a proxy since firm investments take frequently zero values. I will estimate the production function presented in equation 3 in the LP spirit, and proceed very similar to ACF.[9] The identification strategy of the production function parameters is briefly presented in the following. In the first stage a scalar observable is used to control for the unobserved productivity. As auxiliary variable the flexible input factor intermediate products is used, which is supposed to be generated as a function of capital and labor input as well as the unobserved productivity, expressed by $x^m_{nt} = \tilde{f}_t(x^k_{nt}, x^l_{nt}, \omega_{nt})$. The key assumption in the first step is the assumption of strict monotonicity of x^m_{nt} in ω_{nt}. This assumptions implies invertibility of $\tilde{f}_t$ in ω_{nt}, yielding $\omega_{nt} = \tilde{f}_t^{-1}(x^k_{nt}, x^l_{nt}, x^m_{nt})$, which is then substituted into equation (3) to obtain

$$\begin{aligned} y_{nt} &= \alpha_0 + \sum_i \alpha_i x^i_{nt} + \frac{1}{2}\sum_{ij} \alpha_{ij} x^i_{nt} x^j_{nt} + \tilde{f}_t^{-1}(x^k_{nt}, x^l_{nt}, x^m_{nt}) + \epsilon_{nt} && (4) \\ &= \Phi(x^k_{nt}, x^l_{nt}, x^m_{nt}) + \epsilon_{nt}, && (5) \end{aligned}$$

where $\tilde{f}_t^{-1}$ is treated nonparametrically and

$$\Phi(x^k_{nt}, x^l_{nt}, x^m_{nt}) \equiv \alpha_0 + \sum_i \alpha_i x^i_{nt} + \frac{1}{2}\sum_{ij} \alpha_{ij} x^i_{nt} x^j_{nt} + \tilde{f}_t^{-1}(x^k_{nt}, x^l_{nt}, x^m_{nt}). \tag{6}$$

The specification of the (nonparametric) first step yields a conditional mean given by

$$E\left(y_{nt} | x^k_{nt}, x^l_{nt}, x^m_{nt}\right) = \Phi(x^k_{nt}, x^l_{nt}, x^m_{nt}) \tag{7}$$

with the corresponding moment conditions given by

$$E\left(\epsilon_{nt} | x^k_{nt}, x^l_{nt}, x^m_{nt}, x^k_{n,t-1}, x^l_{n,t-1}, x^m_{n,t-1}, \ldots, x^k_{n1}, x^l_{n1}, x^m_{n1}\right) = 0. \tag{8}$$

In this first stage the parameters of interests, α_0 and $\alpha = (\alpha_i,\ \alpha_{ij})$, are not identified. Instead, a nonparametric estimate of $\Phi(x^k_{nt}, x^l_{nt}, x^m_{nt})$ denoted by $\hat{\Phi}(x^k_{nt}, x^l_{nt}, x^m_{nt})$ is obtained, which will be of further importance. In the second stage the parameters in α, (but not the constant α_0), are identified. To accomplish the identification, the second key assumption lies on the productivity shock ω_{nt}, which is assumed to be a first order Markov process, i.e.,

$$\omega_{nt} = E\left(\omega_{nt} | \omega_{n,t-1}\right) + \xi_{nt}, \tag{9}$$

where ξ_{nt} is an iid error term with $E(\xi_{nt} | \omega_{n,t-1}) = 0$. From equation (6) it follows that

$$\widehat{\alpha_0 + \omega}_{nt}(\alpha) = \hat{\Phi}(x^k_{nt}, x^l_{nt}, x^m_{nt}) - \sum_i \alpha_i x^i_{nt} + \frac{1}{2}\sum_{ij} \alpha_{ij} x^i_{nt} x^j_{nt} \tag{10}$$

the innovations in ω_{nt}, namely $\hat{\xi}_{nt}$, can be estimated by regressing $\widehat{\alpha_0 + \omega}_{nt}(\alpha)$ on a higher order polynomial of $\widehat{\alpha_0 + \omega}_{n,t-1}(\alpha)$ for some initial values for the parameters in α.[10] Finally, α can be estimated by a search over the space of the parameters in α, imposing the moment conditions[11]

$$E\left[\hat{\xi}_{nt}(\alpha)\begin{pmatrix} x^k_{nt} \\ x^l_{n,t-1} \\ x^m_{n,t-1} \\ (x^k_{nt})^2 \\ (x^l_{n,t-1})^2 \\ (x^m_{nt-1})^2 \\ x^l_{n,t-1}x^k_{nt} \\ x^m_{n,t-1}x^k_{nt} \\ x^m_{n,t-1}x^l_{n,t-1} \end{pmatrix}\right] = 0. \tag{11}$$

[9] Also see De Loecker and Warzynski (2012) for a further application.

[10] As initial values I use the estimated coefficients of an OLS regression of y_{nt} on all variables of the gross output production function.

[11] The choice of the instruments in the moment equation (11) is related to the timing assumption mentioned above. Since I suppose that firms chose their capital input at $t-1$, whereas the flexible input factor labor is supposed to be chosen at t, I use the instruments x^k_{nt} and $x^l_{n,t-1}$ that should be orthogonal to the shocks in innovation, given by ξ_{nt}.

Note that the moment conditions are derived from the first order Markov assumption (given in equation (9)), implying orthogonality between the production input factors and the innovation to productivity, ξ_{nt}.

4.1.1 Productivity index

By obtaining estimates of the production function parameters I construct for each firm a productivity index following Pavcnik (2002). In most cases, firms productivity can be seen as a residual between the observed and the predicted (conditional mean) value. The proposed productivity index, instead, is constructed as the difference between individuals firms productivity and that of a reference firm, where the reference firm is an artificial (mean) firm for a specific year. I choose 2014 as the reference year. As Pavcnik (2002) (and the concerned literature therein) argues, constructing the productivity index in this way has the desired properties of transitivity and insensitivity to units of measurement. More precisely, firms' productivity index, denoted by $\tilde{\omega}_{nt}$, is normalized as follows

$$\tilde{\omega}_{nt} = y_{nt} - \sum_i \hat{\alpha}_i x^i_{nt} + \frac{1}{2}\sum_{ij} \hat{\alpha}_{ij} x^i_{nt} x^j_{nt} - (\bar{y}_r - \hat{y}_r)\,, \tag{12}$$

where $\bar{y}_r = N_r^{-1} \sum_{n=1}^{N_r} y_{nr}$, with the reference year $r = 2014$ and N_r the number of firms in this year. The predicted log output of the reference firm is given by

$$\hat{y}_r = \sum_i \hat{\alpha}_i \bar{x}^i_r + \frac{1}{2}\sum_{ij} \hat{\alpha}_{ij} \overline{x^i_r x^j_r}, \tag{13}$$

where $\overline{x^i_r x^j_r}$ is the average of the product of x^i_r and x^j_r.

4.1.2 Output elasticities

The estimation of the technology parameters of the translog production function now also allows us to obtain for each firm (and time period) output elasticities as well as economies of scale estimates. Output elasticities for each input factor are obtained by

$$\widehat{\text{elas}}^i_{nt} = \frac{\partial y_{nt}}{\partial x^i_{nt}} = \hat{\alpha}_i + \sum_j \hat{\alpha}_{ij} x^j_{nt}. \tag{14}$$

The firm-level estimate for returns to scale is given by

$$\widehat{\text{scale}}_{nt} = \widehat{\text{elas}}^k_{nt} + \widehat{\text{elas}}^l_{nt} + \widehat{\text{elas}}^m_{nt}. \tag{15}$$

4.2 Productivity persistence

In order to analyze the distributional dynamics of firms' productivity I follow Johnson (2000) and Johnson (2005), who investigates distributional dynamics in the context of U.S. cross-country production and income. Let the productivity distribution at t is given by $f_t(\omega)$. This distribution evolves over time and takes the form $f_{t+\tau}$ at $t+\tau$ for $\tau > 0$. Assuming that a time invariant and first-order Markov process drives the evolution of the productivity distribution, the distribution at $t+\tau$ can be described by $f_{t+\tau}(\tilde{\omega}) = \int_0^\infty g_\tau(\tilde{\omega}|\omega) f_t(\omega) d\omega$, where $g_\tau(\tilde{\omega}|\omega)$ denotes the productivity distribution at $t+\tau$ conditioned on firms' productivity at t and $f_t(\omega)$ the marginal productivity distribution in t.[12] In order to analyse how firms improve their productivity, conditional on their productivity level in the past, I am interested in the estimation of $g_\tau(\tilde{\omega}|\omega)$, by

$$\hat{g}_\tau(\tilde{\omega}|\omega) = \frac{\hat{\phi}_{t-\tau,t}(\omega,\tilde{\omega})}{\hat{f}_{t-\tau}(\omega)}, \tag{16}$$

with the joint and marginal distribution $\hat{\phi}_{t-\tau,t}(\omega,\tilde{\omega})$ and $\hat{f}_{t-\tau}(\omega)$. These objects are estimated based on the couples $\{\tilde{\omega}_{nt}, \omega_{n,t-\tau}\}_{n=1,t=1}^{N,T_\tau}$, using conventional nonparametric kernel density esti-

[12] $g_\tau(\tilde{\omega}|\omega)$ that can be seen as a representation of continuous state transitions, whereas in the case of discrete state measures transition matrices are often employed for the representation.

mation methods (Li and Racine, 2011, Chapter 1 and 5), given

$$\hat{\phi}_{t-\tau,t}(\omega,\tilde{\omega}) = \frac{1}{NT_\tau h_\omega h_{\tilde{\omega}}} \sum_{n=1}^{N} \sum_{t=1}^{T_\tau} K\left(\frac{\omega_{n,t-\tau}-\omega}{h_\omega}\right) K\left(\frac{\tilde{\omega}_{nt}-\tilde{\omega}}{h_{\tilde{\omega}}}\right) \tag{17}$$

$$\hat{f}_{t-\tau}(\omega) = \frac{1}{NT_\tau h_\omega} \sum_{n=1}^{N} \sum_{t=1}^{T_\tau} K\left(\frac{\omega_{n,t-\tau}-\omega}{h_\omega}\right), \tag{18}$$

where $K(.)$ denotes a second order Gaussian kernel and h_ω and $h_{\tilde{\omega}}$ are the bandwidths, optimally chosen by the data-driven likelihood cross-validation method.[13]

4.3 Productivity decomposition

In the following I present the adopted methodology to measure aggregate productivity growth and its decomposition into those parts contributed by surviving, entering and exiting firms.

4.3.1 The Dynamic Olley-Pakes Decomposition

Olley and Pakes (1996) presented static approach (without entry and exit) to calculate an industry's aggregate productivity. The basic expression of this measure was given in equation (1), which is simply, for a given industry, a weighted average of firms' productivity, weighted by firms' market shares, s_{nt}. It can be shown that this weighted average can be further separated by

$$\begin{aligned}\Omega_t &= \bar{\omega}_t + \sum_n (s_{nt} - \bar{s}_t)(\omega_{nt} - \bar{\omega}_t) \\ &= \bar{\omega}_t + N_t cov(s_{nt}, \omega_{nt})\end{aligned} \tag{19}$$

where $\bar{\omega}_t = N_t^{-1} \sum_{n=1}^{N_t} \omega_{nt}$ denotes an industry's unweighted average productivity, N_t, the number of active firms and $\bar{s}_t = 1/N_t$ the mean market share. That is, an industry's aggregate productivity is composed of an unweighted average of firms' productivity and the covariance between firms' productivity and their market shares. Note that, since the cov operator already contains division by N_t, we need to premultiply $cov(.)$ by N_t, such that the second equality is equivalent to the first equality. Considering the change in aggregate productivity, $\Delta\Omega_t = \Delta\bar{\omega}_t + \Delta N_t cov(s_{nt}, \omega_{nt})$, it can be seen that aggregate productivity growth is transmitted by the change in the unweighted mean (referred to within-change) and by a change in the covariance between firms' productivity and market shares (referred to between-change).

In a dynamic setting with entry and exit, Ω_t is further composed of the aggregate productivity of surviving, entering and exiting firms. To measure the contribution of the three firms groups I adopt the Dynamic Olley-Pakes Decomposition (DOPD, henceforth; (Melitz and Polanec, 2015)).[14] Let $S_{Gt} = \sum_{n \in G} s_{nt}$ be the aggregate market share of a group G, where $G \in \{E, S, X\}$, represents either the set of entrants, survivors and exitors. A group's aggregate productivity is then defined by $\Omega_{Gt} = \sum_{n \in G} (s_{nt}/S_{Gt})\, \omega_{nt}$. Considering two periods, the aggregate productivity can be expressed as a weighted mean of the mentioned groups:

$$\Omega_1 = S_{S1}\Omega_{S1} + S_{X1}\Omega_{X1} = \Omega_{S1} + S_{X1}(\Omega_{X1} - \Omega_{S1}) \tag{20}$$

$$\Omega_2 = S_{S2}\Omega_{S2} + S_{E2}\Omega_{E2} = \Omega_{S2} + S_{E2}(\Omega_{E2} - \Omega_{S2}). \tag{21}$$

In each period the aggregate productivity is a weighted average of the aggregate productivity measures of the underlying firm groups: More precisely, at $t = 1$ the overall aggregate productivity is composed of the aggregate productivity of surviving firms from $t = 1$ to $t = 2$ as well as the aggregate productivity of exiting firms (those firms active for the last time in $t = 1$). Instead, at $t = 2$, the overall aggregate productivity is composed of the aggregate productivity of surviving firms and entering firms. The difference in aggregate productivity between both periods can be derived by $\Delta\Omega = \Omega_2 - \Omega_1$ and is given by

$$\begin{aligned}\Delta\Omega &= \underbrace{(\Omega_{S2} - \Omega_{S1})}_{\text{Survivors}} + \underbrace{S_{E2}(\Omega_{E2} - \Omega_{S2})}_{\text{Entrants}} + \underbrace{S_{X1}(\Omega_{S1} - \Omega_{X1})}_{\text{Exitors}} \\ &= \Delta\bar{\omega}_S + \Delta N_S cov_S + S_{E2}(\Omega_{E2} - \Omega_{S2}) + S_{X1}(\Omega_{S1} - \Omega_{X1}).\end{aligned} \tag{22}$$

[13] For the empirical application I use the **np** package from R (Hayfield and Racine, 2015). Optimal bandwidths are computed using the the likelihood cross-validation method, which is, compared to the least square cross-validation method, computationally more efficiency, see Henderson and Parmeter (2015, p. 67).

[14] See Autor et al. (2019) for an application of the DOPD approach with respect to labor share dynamics in U.S. manufacturing industries.

As already pointed out by Griliches and Regev (1995), entering and exiting firms can have a positive or negative contribution to aggregate productivity depending on the considered reference level of productivity. As can be seen in equation (22), in the DOPD approach entering firms' contribution to aggregate productivity increases with their aggregate productivity level in the second period, given by Ω_{E2}. That is, only if the aggregate productivity of entering firms is higher compared to the aggregate productivity of surviving firms, entering firms contribute positively to aggregate productivity growth. Similarly, exiting firms' only contribute positively to aggregate productivity if their aggregate productivity in the first period, given by Ω_{X1}, is lower compared to the aggregate productivity of surviving firms in the same period. Finally, surviving firms only contribute positively to aggregate productivity, if their aggregate productivity in the second period is higher compared to the first period.

4.3.2 Identification of firm entry and exit for the DOPD analysis

Generally, it would be desirable to present the yearly changes throughout the whole sample period. However, since my data set comprises 23 years I only report aggregate productivity changes in three-year waves. In Section 5.3, I report productivity changes for the waves 1998, 2001, 2004, 2007, 2010, 2013 and 2016, where all changes are relative to the initial measure in 1995.[15]

Since the DOPD approach aims to assign the contribution of surviving, entering and exiting firms to aggregate productivity change I need to identify firms that have survived, entered or exited between the initial year 1995 and the given wave. This means that I need to define firms' status of either survivor, entrant or exitor for the case of longer time spans. For this purpose I make use the definitions made in Section 3.2 by adding some extra constraints: Let t_1 be the initial year (here 1995) and $t_2 \in \{1998, 2001, 2004, 2007, 2010, 2013, 2016\}$ be the last year of the $[t_1, t_2]$ wave. A firm is said to survive between t_1 and t_2 if the firm is active both in t_1 and t_2. Moreover, a firm is said to have entered the market between t_1 and t_2 if (i) the firm has been registered to have entered the market, i.e., $e^+_{nt} = 1$ for some t with $t_1 < t \leq t_2$, and (ii) if the firm was inactive in t_1 but active in t_2, i.e. $a_{n,t_1} = 0$ and $a_{n,t_2} = 1$. Analogously, a firm is said to have exited the market between t_1 and t_2 if (i) the firm has been registered to have exited the market, i.e., $e^-_{nt} = 1$ for some t with $t_1 \leq t < t_2$ and if (ii) the firm was active in t_1 but inactive in t_2, i.e., $a_{n,t_1} = 1$ and $a_{n,t_2} = 0$.

4.4 Productivity differences between firm groups

I am interested in investigating productivity differences between different firm groups, such as entering, surviving and exiting firms but also between small, medium and big sized firms. For this purpose follow Fariñas and Ruano (2005) who analyzed these aspects for Spanish manufacturing industries. The analysis is conducted in two parts: (i) by a graphically comparison between the empirical cumulative density function (ECDF) of the firms belonging to the different groups and (ii) by statistically testing differences among these distributions.

(i) Graphical comparison

In order to graphically analyze the distributions between different groups of firms I visualize the CDF's of the corresponding firm group. This allows to compare the whole productivity distributions of different groups of firms, instead of only comparing single moments, such as the mean or median.

Let $\hat{F}_G(c)$ be the productivity ECDF of a specific firm group, where

$$\hat{F}_G(c) = \frac{1}{N_G} \sum_{n \in G} \mathbf{1}_{[\hat{\omega}_n \leq c]}, \tag{23}$$

where $\mathbf{1}_{[A]}$ denotes a dummy variable equal to 1 if the condition A in brackets is satisfied and 0 otherwise. The intuition of the concept of (first-order) stochastic dominance is, if the position of productivity ECDF of group one is consistently located to the right of the ECDF of group two, then the distribution of group two stochastically dominates the distribution of group one. For each percentile, firms' productivity levels belonging to group two are higher compared to group one.

(ii) Testing procedure

Let F_1 and F_2 be the CDF's of firm productivity of two groups of firms (such as entrants and survivors or exitors and survivors), for a given period t. First order stochastic dominance of F_1

[15]See Appendix D providing the DOPD analysis with 2007 as initial year and aggregate productivity growth rates reported for the waves 2010, 2013 and 2016.

with respect to F_2 implies $F_1(\omega) - F_2(\omega) \leq 0$, with strict inequality for a specific productivity level ω, where $P(\omega \in \mathbb{R}) = 1$. The Kolmogorov-Smirnov test allows to test for stochastic dominance.[16] First, the two-sided test allows to test whether the distributions F_1 and F_2 follow the same law and is given by

$$H_0 : \sup_{\omega \in \mathbb{R}} |F_1(\omega) - F_2(\omega)| = 0 \quad \text{vs.} \quad H_A : \sup_{\omega \in \mathbb{R}} |F_1(\omega) - F_2(\omega)| \neq 0, \tag{24}$$

The one-sided test, allows to specifically test which of the two distributions (first order) stochastically dominates the other and is given by

$$H_0 : \sup_{\omega \in \mathbb{R}} \{F_1(\omega) - F_2(\omega)\} = 0 \quad \text{vs.} \quad H_A : \sup_{\omega \in \mathbb{R}} \{F_1(\omega) - F_2(\omega)\} > 0. \tag{25}$$

The respective test statistics for the two- and one-side test are given by

$$\text{KS}_N^{\text{two}} = \sqrt{\frac{N_1 \cdot N_2}{N}} \sup_{\omega \in \mathbb{R}} |T_N(\omega)| \quad \text{and} \quad \text{KS}_N^{\text{one}} = \sqrt{\frac{N_1 \cdot N_2}{N}} \sup_{\omega \in \mathbb{R}} T_N(\omega), \tag{26}$$

where $T_N(\omega) = \hat{F}_{1,N_1}(\omega) - \hat{F}_{2,N_2}(\omega)$, with $\hat{F}_{1,N_1}$ and $\hat{F}_{2,N_2}$ the empirical CDF's of F_1 and F_2 and $N = N_1 + N_2$ denotes the total number of observations from both distributions. Note that for many situations it is possible that the one- and the two-sided test statistics take the same value, with equal or varying p-values. Furthermore, the Kolmogorov-Smirnov test relies on the assumption of independence of the observations. Firms productivity shocks, however, are supposed to be first-order Markov and, therefore, assumed to be temporally dependent. By this reason I conduct the test for chosen cohorts of entry and exit. That means, we compare productivity distributions between entrants and survivors as well as between exitors and survivors, for some chosen years, and rely only on independence over firms, and not over time periods.

5 Empirical results

Section 5.1 presents for each considered 2-digit industry the estimates of the production function as well as the corresponding output elasticities and returns to scale estimates; Section 5.2 presents the results of the analysis of productivity persistence; Section 5.3 reports results of the Dynamic Olley-Pakes Productivity decomposition; and, finally, Section 5.4 discusses the findings concerning productivity differences between different firm groups.

5.1 Production function estimates

In this section I will present the results of the estimates of the translog gross output production function. In Appendix B Table 22 illustrates for each considered two-digit industry the parameter estimates of the translog production function. However, I mainly discuss the estimates of firms' output elasticities and returns to scale, given in Table 4.[17] Output elasticities and returns to scale are estimated on the firm level, i.e. for each firm and year. The reported statistics, however, are across all years (1994-2016) for a given two-digit industry. Generally, if a firm's output elasticity (with respect to capital, labor and/or materials) is estimated to be negative, the monotonicity condition is not fulfilled. That is, an increase in the use of a production factor leads to a decrease in the output level. This could be due to miss-specification in the underlying production function and/or measurement errors in the data.

Considering first the output elasticities with respect to the input factor capital. It can be seen that the industry for "pulp and paper" (17) reveals the highest output elasticity with respect to the factor capital of 0.21, i.e. at the median an increase in 1% of capital is associated with an increase of 0.21% of production. Generally, the median and the mean values are relatively close, indicating for a low degree of skewness with respect to the distribution of firms' output elasticity with respect to capital. The last column reports the correlation between firms' output level and the output elasticities. It can be seen that especially for the industry "food processing" (10) there is relatively strong positive correlation, i.e. firm with higher output levels also tend to have a higher output elasticity with respect to capital. Many other industries, however, to show the opposite sign, such as the industries "pulp and paper" (17), "chemical products" (20), "basic metals" (24), "automobiles" (29) and "furniture" (31). Considering the estimates of output

[16] See Kolmogorov (1933) and Smirnov (1939).
[17] Appendix B also provides exemplary R code of the estimation procedure.

elasticity with respect to the input factor labor substantial differences among industries can be seen. Here, the industry "food processing" (10) turns out to have the lowest median elasticity, given by 0.33. Instead, the industry "machines" (28) reveals the highest median elasticity, give by 0.63. Generally, however, firms' output elasticity with respect to labor is estimated to be higher than for capital. Moreover, the correlation between firms' log production and output elasticity here also shows varying signs among industries. More precisely, while I measure a positive correlation for the industries "food", "wood", "metals" and "furniture", a negative correlation is found for the rest of the industries. The output elasticity with respect to the remaining input factor materials show the highest (median) amplitude. Here the lowest (highest) median output elasticity is measured for the industry "food" (wood), given by 0.42 (0.61). Contrarily to the correlation between production and the elasticities with respect to capital and labor, the correlation between production and the output elasticity with respect to materials is positive throughout all industries. That is, the higher a firm's production level, the more efficiently materials are transformed into output. Finally, Table 4 reports at the bottom the resulting returns to scale. Considering again the median value, only the industries "food" shows considerable decreasing returns to scale at the median, given by 0.88. That is, an increase in inputs by 1% is associated with an increase in production by only 0.88%. The other industries reveal slightly increasing returns to scale, where the highest median is estimated for the industry "chemical products", given by 1.13. For this industry and increase in inputs of 1% is associated with an increase in production by 1.131%. Furthermore, except for the industries "pulp and paper", "chemical products" and "basic metals", I measure a positive correlation between firms' production level and the degree of returns to scale.

Table 4: Output elasticities and returns to scale

Output elasticity: capital						
Industry	min	median	mean	max	sd	$cor\left(y_{nt}, \widehat{\text{elas}}_{nt}^{k}\right)$
10 food	−0.97	0.15	0.16	1.52	0.12	0.37
13 textiles	−0.18	0.10	0.10	0.42	0.05	−0.35
16 wood	−0.35	−0.00	−0.01	0.28	0.06	0.07
17 pulp/paper	−0.32	0.21	0.21	0.63	0.07	−0.35
20 chemical prod.	0.00	0.20	0.20	0.50	0.04	−0.49
24 metals	−0.29	0.10	0.10	0.48	0.05	0.06
28 machines	−0.38	0.08	0.09	0.87	0.10	−0.70
29 automobiles	−0.60	0.06	0.06	1.20	0.10	−0.31
31 furniture	−0.42	0.08	0.08	0.56	0.06	−0.34
Output elasticity: labor						
Industry	min	median	mean	max	sd	$cor\left(y_{nt}, \widehat{\text{elas}}_{nt}^{l}\right)$
10 food	−0.53	0.33	0.34	1.51	0.12	0.04
13 textiles	−0.17	0.47	0.48	1.16	0.10	−0.40
16 wood	−0.23	0.44	0.44	1.00	0.10	0.43
17 pulp/paper	−0.07	0.42	0.43	1.23	0.12	−0.43
20 chemical prod.	−0.23	0.38	0.38	1.13	0.09	−0.62
24 metals	−0.24	0.34	0.34	0.80	0.14	−0.87
28 machines	−0.12	0.63	0.63	1.53	0.13	0.25
29 automobiles	−0.73	0.52	0.52	1.81	0.20	−0.51
31 furniture	−0.27	0.37	0.38	1.18	0.14	0.55
Output elasticity: materials						
Industry	min	median	mean	max	sd	$cor\left(y_{nt}, \widehat{\text{elas}}_{nt}^{m}\right)$
10 food	−0.45	0.42	0.41	1.34	0.10	0.35
13 textiles	−0.11	0.48	0.48	0.92	0.09	0.66
16 wood	0.54	0.61	0.61	0.67	0.02	0.71
17 pulp/paper	−0.08	0.45	0.45	1.02	0.10	0.59
20 chemical prod.	0.23	0.54	0.55	0.86	0.05	0.75
24 metals	0.21	0.60	0.60	1.01	0.12	0.93
28 machines	−0.12	0.44	0.44	0.76	0.06	0.64
29 automobiles	−0.33	0.49	0.51	1.38	0.19	0.81
31 furniture	−0.22	0.53	0.53	1.17	0.08	0.06
Returns to scale						
Industry	min	median	mean	max	sd	$cor\left(y_{nt}, \widehat{\text{scale}}_{nt}\right)$
10 food	0.35	0.88	0.91	1.75	0.12	0.67
13 textiles	0.90	1.06	1.06	1.26	0.03	0.03
16 wood	0.60	1.04	1.04	1.37	0.08	0.77
17 pulp/paper	0.80	1.08	1.09	1.38	0.05	−0.39
20 chemical prod.	0.96	1.13	1.13	1.38	0.04	−0.76
24 metals	0.67	1.03	1.04	1.48	0.06	−0.15
28 machines	0.75	1.15	1.15	1.65	0.07	0.01
29 automobiles	0.50	1.09	1.09	1.80	0.08	0.23
31 furniture	0.62	0.97	0.98	1.38	0.08	0.74

5.2 Productivity persistence

I now turn to the investigation of productivity dynamics based on the firm-level productivity estimated through the translog production function. My first interest is to study persistence in firms' level of productivity, or, in other words firms' ability to improve their productivity throughout time. As outlined in Section 4.2 I estimate the density of firms productivity in t conditional on their productivity at $t-1$, that is $\tau = 1$. Figure 4 depicts the conditional density in a 3D plot and Figure 5 the corresponding contour plot. While the 3D plot rather provides a general impression of the conditional density, the contour plot allows for a precise analysis. Generally, both Figures show that (measured over the whole period) from one year to another, firms' productivity remains for all productivity levels basically the same. The 3D plot illustrates this by a very peaked density along the diagonal. Considering the contour plot, the bisecting line represents those productivity levels for which firms' productivity at t exactly hits the productivity at $t-1$. It can seen that the highest density (represented by the inner contour lines) throughout the whole support of the productivity measures is very closely located around the bisecting line. This means that productivity is highly persistent with respect to all levels of productivity. These findings are similar to Foster et al. (2006), analysing the U.S. retail sector and Polanec (2004), for Slovenian manufacturing industries. Also see Appendix C for conditional density estimates setting $\tau = 3$, for which case very similar patterns are shown.

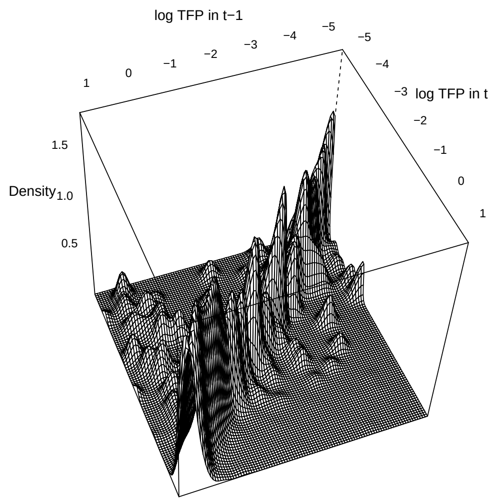

Figure 4: 3D plot of the conditional density $\hat{g}_1(\hat{\tilde{\omega}}|\hat{\omega})$

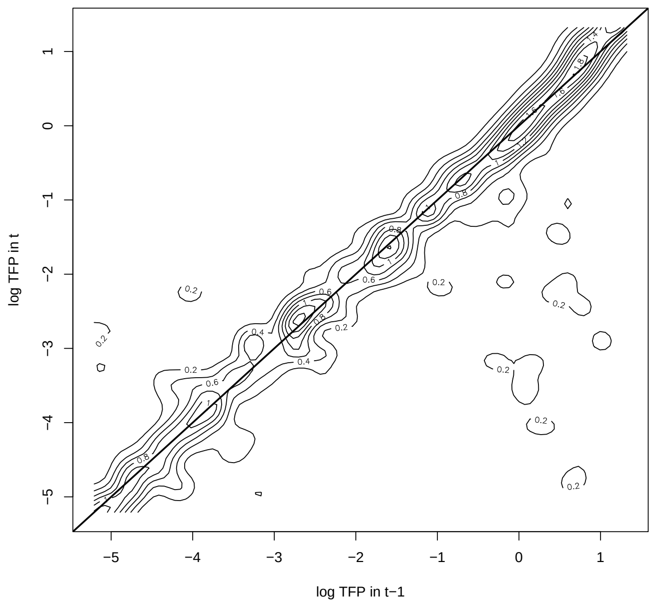

Figure 5: Contour of the conditional density $\hat{g}_1(\hat{\tilde{\omega}}|\hat{\omega})$

In Section 4.1, equation (9) I assume that firms' productivity follows a first order Markov process. Specifically,

$$\begin{aligned}\omega_{nt} =& E(\omega_{nt}|\omega_{n,t-1}) + \xi_{nt} \\ =& \varphi(\omega_{n,t-1}) + \xi_{nt},\end{aligned}$$

where $E(\omega_{nt}|\omega_{n,t-1}) \equiv \varphi(\omega_{n,t-1})$ with the assumption $E(\xi_{nt}|\omega_{n,t-1}) = 0$. I estimate nonparametrically the conditional mean function, φ, applying the local-linear least squares estimator, which also enables to recover an estimate for the gradient $\varphi'(\omega_{n,t-1}) = \partial\varphi(\omega_{n,t-1})/\partial\omega_{n,t-1}$. Note that, in total, I have 997,764 estimates of the couple $(\hat{\omega}_{nt}, \hat{\omega}_{n,t-1})$ over all industries and years. However, since nonparametric estimation is computationally burdensome I randomly select only 20,000 couples to visualize the results. Figure 6 shows the estimated conditional mean function, here represented by the solid line. Except at the boundaries, the function is shown to be increasing and almost linear in $\hat{\omega}_{n,t-1}$, which means that firms with higher productivity in the past period are very likely to be higher productive in the current period. The (bootstrapped) point-wise 95 % confidence interval is represented by the dashed line. Generally, the estimated function always lies within the confidence intervals, indicating relatively reliable nonparametric estimates. This is especially the case for the region $\hat{\omega}_{n,t-1} \in [-5, 5]$, whereas at the boundaries, where less observations are available, the confidence intervals become wider and, thus, estimates are less precise. Furthermore, the gradient estimate is shown in Figure 7. The estimated gradient, represented by the solid line, shows to be relatively constant and positive within the region $\hat{\omega}_{n,t-1} \in [-5, 5]$. This indicates that there are no substantial differences between low and high productivity firms with respect to the importance of firms past productivity level on their current productivity level. Note that the average marginal effect of $\hat{\omega}_{n,t-1}$ on $\hat{\omega}_{nt}$ corresponds to the mean of all point estimates of the gradient $\hat{\varphi}'(\hat{\omega}_{n,t-1})$. For this sample the estimated average marginal effect is given by 0.77, which underlines strong persistence in firm level productivity.

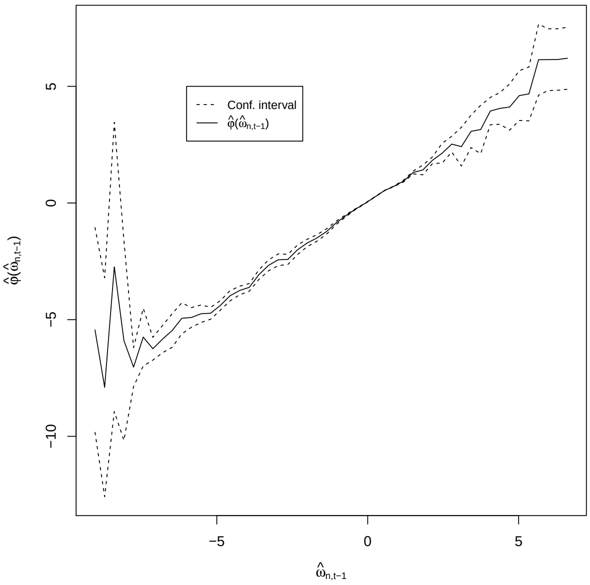

Figure 6: Nonparametric estimate of the conditional mean function $\hat{\varphi}(\hat{\omega}_{n,t-1})$.

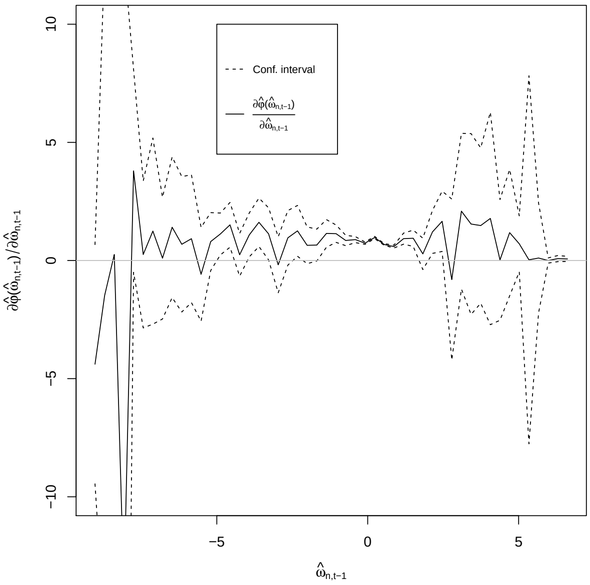

Figure 7: Nonparametric estimate of the gradient $\partial\hat{\varphi}(\hat{\omega}_{n,t-1})/\partial\hat{\omega}_{n,t-1}$.

5.3 Aggregate productivity change

I now turn to the results of aggregate productivity change with respect to different industries. Table 5 reports for all industries the aggregate productivity growth rates, decomposed along the Dynamic Olley-Pakes Productivity Decomposition (Melitz and Polanec, 2015). That is, into the contribution to aggregate productivity growth of surviving, entering and and exiting firms. Table 5 reports, for each industry separately, the decomposition given in equation (22). To sustain the interpretation of the results I also rely on Tables 6 and 7, reporting the aggregate measures (aggregate market shares and productivity) for the initial year 1995 (t_1) and those measures for the considered waves (t_2). That is, Tables 6 and 7 correspond to equations (20) and (21), respectively.

5.3.1 Industries' total aggregate productivity growth

Beginning with the industries' total aggregate productivity growth. A positive growth between the base year 1995 (t_1) and $t_2 = 1998, \ldots, 2016$ occurs if the overall productivity - composed of the aggregate of survivors, entrants and exitors - is higher at t_2 compared to t_1, i.e. $(\Omega_2 - \Omega_1) > 0$. In Table 5, column(s) "total change", it can be seen that in 2016 most of the industries have experienced an increase in aggregate productivity, relative to 1995. More precisely, the growth rates are given by -19.49% (food), 33.38% textiles, 40.33% (wood), 22.34% (pulp/paper), 15.97% (chemical prod.), 24.66% (metals), 25.11% (machines), -30.40% (automobiles) and 21.66% furniture.

5.3.2 Contribution of survivors to aggregate productivity growth

As outlined in Section 4.3, the group of survivors' contribution to aggregate productivity is measured by the within (unweighted) firm productivity change and between contribution (reallocation effect of market shares among firms). Here survivors only contribute positively to aggregate productivity change if their aggregate productivity in the considered wave, $t_2 = 1998, \ldots, 2016$, is higher compared to the same groups' aggregate productivity in the base year 1995 t_1.

Table 5 shows that the aggregate productivity of surviving firms has improved over time. It can be seen that the groups' within productivity change drives in many industries the aggregate productivity improvement, whereas the between change seems to be less important.[18] More precisely, except for the industry "metals" in 2007, I measure positive growth rates with respect to surviving firms' within change, ranging (in 2016) between 15.80% (food) and 37.55% (wood). For many industries and years I measure a negative sign in the between change, i.e. a negative growth induced by a negative change in the correlation between firms' productivity level and their market shares. Especially for the industries "food", "chemical products" and "machines", I measure a very strong negative contribution to aggregate productivity growth through reallocation effects. In these cases, surviving firms' aggregate between change compensates a large part (and sometimes even exceeds) the within change. Only for the "textile" industry I measure a positive, albeit slight, contribution to aggregate productivity growth through reallocation effects. Haltiwanger (2011), refers to "dynamic allocative efficiency" if the change in the covariance between firm productivity and market share increases over time. A positive sign implies that market shares wander from low to higher productivity firms, generating a higher consumer and producer surplus. If this is the case Haltiwanger (2011) speaks about a "well-functioning economy". My measures indicate that, in this sense, many of the considered French manufacturing industries are not well-functioning. The general impression of a higher importance if the within-growth component, compared to the between growth component, confirms findings shown by Griliches and Regev (1995) and Melitz and Polanec (2015) for the Israeli and Slovenian manufacturing industries, respectively. However, my findings of negative contribution to aggregate productivity growth by the between-growth component contradicts finding in Melitz and Polanec (2015) for the Slovenian case. Also the general picture in Pavcnik (2002) for Chilean manufacturing industries shows a positive contribution to aggregate productivity growth by the between-growth component.

5.3.3 Contribution of entrants to aggregate productivity growth

I turn now to the contribution of firm entry to aggregate productivity change, shown in Table 5 column(s) "contribution entrants". Recall that a positive contribution of the group of entrants to aggregate productivity growth is measured if the aggregate productivity of those firms that entered between the base year 1995 (t_1) and the considered wave ($t_2 = 1998, 2001, \ldots, 2016$) is higher in t_2 compared to the the aggregate productivity of those firms that survived between the initial year and t_2, i.e. $(\Omega_{E2} - \Omega_{S2}) > 0$. If this is the case, the group of entrants has a relatively high level of aggregate productivity, compared to survivors and, hence, the industry is enriched by well preforming new firms. For many cases (industries and years) the figures show that firm entry has contributed positively to the overall aggregate productivity change. For instance, for the industry "food" entering firms contribute positively to aggregate productivity change between 5.05% in 2001 to 9.3% in 2004, with respect to 1995. Only in 2016, firms that entered with respect to 1995 turn out to contribute negatively to aggregate productivity growth. Similar results are found for other industries such as "wood", "chemical products", "metals", "machines" and "furniture". The industry "automobiles", however, represents a counter example, where I measure a massive negative contribution of entering firms to aggregate productivity change. For this industry, firms

[18] Very low between changes, are also reported Bartelsman et al. (2013) for other countries.

that entered between 1995 and 2016 contribute negatively by -85.97% to the overall change, which is a much higher contribution than that of the group of surviving firms, given by 19.39%. That is, according to my estimates, the group of entrants in the French automobile industry exhibits a considerably lower level of aggregate productivity compared to the group surviving firms.

Table 7 illustrates market shares and aggregate productivity of both surviving and entering firms, measured in t_2, i.e. for each of the reported industries and waves. Considering the automobile industry, it can be seen that the aggregate productivity of entrants (column(s) Ω_E) is considerably lower compared to the aggregate productivity of survivors (column(s) (Ω_S), which consequently induces a negative contribution of entrants. Furthermore, for the automobile industry a sharp decrease in the market share of surviving firms is observed (column S_S) from 2001 on, and, analogously, a sharp increase in the aggregate market shares of the group of entering firms (column S_E). Hence, the negative difference between the aggregate productivity of entering and surviving firms becomes more and more weighted by the group of entrants increasing aggregated market share, which further increases the negative contribution of entering firms to aggregate productivity growth. For those industries (and years) for which a positive contribution of entrants is reported, it can be seen that the aggregate productivity of entering firms is higher than that of surviving firms. Note that, generally, market shares of survivors (entrants) decrease (increase) the longer the time span between 1995 and t_2 becomes, which is also reflected in the number of observed survivors/entrants.[19] Both Baily et al. (1992) (US data) and Melitz and Polanec (2015) find that the group of entering firms contribute less to aggregate productivity growth compared to the group of surviving firms. Beside some exception, the global pictures of my results confirms these findings for French manufacturing industries.

5.3.4 Contribution of exitors to aggregate productivity growth

Finally, I discuss the contribution of exiting firms to aggregate productivity change, shown in Table 5, columns "contribution exitors". There is a positive contribution of the group of exitors to aggregate productivity growth if the aggregate productivity of those firms that have exited between the base year 1995 (t_1) and the considered wave ($t_2 = 1998, 2001, \dots, 2016$) is lower in t_1 compared to the the aggregate productivity of those firms that survived between the initial year and t_2, i.e. $(\Omega_{S1} - \Omega_{X2}) > 0$. Generally, I find varying signs with respect to the contribution of exiting firms to aggregate productivity growth. For instance, throughout the whole period, the group of exiting firms in the food industry has contributed negatively to aggregate productivity growth, ending up in 2016 with a contribution of -21.58%, whereas survivors and entrants contribute with 15.8% and -4.88%. Also, the strong negative contribution of exiting firms leads for the waves 2013 and 2016 to a negative overall aggregate productivity growth. That is, this industry has lost over time more and more firms that exhibit in 1995 a higher aggregate productivity compared to the group of surviving firms.

The industry for automobiles shows the opposite. Here exitors contribute significantly positively to aggregate productivity change, up to 41.22% in 2016. This strong positive contribution to aggregate productivity growth of both exitors and survivors in the automobile industry, however, does not suffice to compensate the high negative contribution of entering firms. This results implies that, according to my measures, the industry has experienced a drop out of firms that were relatively less productive with respect to those firms that have survived. In other industries such as "textiles", "wood", "pulp/paper", "chemical products" and "machines", exitors contribute positively but less importantly to the total aggregate productivity change. Table 6 reports market shares and aggregate productivity of both surviving and exiting firms. Considering the food industry, where exitors contribute negatively to the aggregate productivity growth, it can be seen that their aggregate productivity (column Ω_X) is always higher than the aggregate productivity of surviving firms (column Ω_S). Instead, the automobile industry, exhibiting a significant positive contribution of exitors, shows a higher aggregate productivity of the group of surviving firms (in 1995). The reported group market shares of survivors and exitors (columns S_S and S_X) are measured for the initial year 1995 of those firms surviving/exiting until t_2. For all industries the same pattern of decreasing (increasing) market shares in 1995 of those firms surviving (exiting) until year t_2.[20] Similar to the case of the contribution of entering firms to aggregate productivity growth Baily et al. (1992) and Melitz and Polanec (2015) find minor importance of the contribution of exiting firms to aggregate productivity growth (compared to the contribution of surviving firms), which goes in line with my findings.

[19] Note that for each given couple, t_1 and t_2, the market shares of surviving and entering firms add up to 100%.
[20] Note that market shares of surviving and exiting firms always add up to 100%.

Table 5: Aggregate Productivity Decomposition (DOPD): 1995 - 2016[a]

Industry	Year	Contribution Survivors		Contribution Entrants	Contribution Exitors	Total Change[b]	Industry	Year	Contribution Survivors		Contribution Entrants	Contribution Exitors	Total Change[b]
		Within	Between						Within	Between			
10	1998	6.03	-9.21	5.93	-5.32	-2.57	24	1998	6.54	-3.44	0.66	-0.78	2.98
Food	2001	2.30	-9.07	5.05	-4.71	-6.43	Metals	2001	6.84	-1.81	17.63	-3.92	18.74
	2004	2.36	-10.30	9.30	-7.28	-5.92		2004	5.22	-2.17	17.75	-4.01	16.79
	2007	6.66	-10.36	8.64	-9.96	-5.02		2007	-8.08	2.59	24.81	-4.06	15.26
	2010	18.72	-12.40	9.25	-16.76	-1.19		2010	1.70	0.56	3.49	-5.04	0.71
	2013	14.49	-11.78	1.07	-19.33	-15.55		2013	15.66	-3.91	-3.77	5.00	12.98
	2016	15.80	-8.83	-4.88	-21.58	-19.49		2016	20.73	-3.71	3.01	4.63	24.66
13	1998	4.96	0.98	0.54	1.26	7.74	28	1998	4.25	-6.50	-0.44	1.54	-1.15
Textile	2001	9.80	1.63	-0.56	1.64	12.51	Machines	2001	10.82	-13.79	3.67	2.69	3.39
	2004	12.18	2.59	-0.98	3.28	17.07		2004	18.69	-13.15	6.95	5.20	17.69
	2007	9.34	6.61	-0.59	3.94	19.30		2007	21.72	-9.89	1.42	6.07	19.32
	2010	19.81	1.55	-0.71	5.01	25.66		2010	30.07	-12.99	-2.72	7.19	21.55
	2013	23.56	2.01	-0.31	5.31	30.57		2013	32.50	-10.21	8.24	-2.70	27.83
	2016	27.73	1.06	-1.27	5.86	33.38		2016	30.62	-5.84	4.98	-4.65	25.11
16	1998	7.16	-5.13	1.26	1.01	4.30	29	1998	7.88	-44.16	2.87	-1.22	-34.63
Wood	2001	14.86	-6.85	-2.71	1.08	6.38	Automobiles	2001	6.37	-0.72	-29.41	39.29	15.53
	2004	22.00	-4.87	1.96	1.66	20.75		2004	8.39	-2.08	-40.38	37.15	3.08
	2007	18.55	-3.37	2.36	1.05	18.59		2007	15.08	-3.63	-42.94	39.16	7.67
	2010	33.39	-7.86	3.82	2.30	31.65		2010	15.17	-5.03	-77.21	37.41	-29.66
	2013	35.74	-8.19	4.98	1.99	34.52		2013	16.68	-5.44	-78.34	38.08	-29.02
	2016	37.55	-6.07	5.42	3.43	40.33		2016	19.39	-5.04	-85.97	41.22	-30.40
17	1998	12.61	-7.89	0.37	1.02	6.11	31	1998	4.05	0.00	1.83	-1.44	4.44
Pulp/	2001	12.40	-5.86	1.61	0.55	8.70	Furniture	2001	7.53	-4.15	4.23	-2.12	5.49
Paper	2004	16.44	-6.88	-0.71	1.60	10.45		2004	13.01	-4.49	4.19	-4.47	8.24
	2007	19.52	-7.19	0.70	0.75	13.78		2007	20.73	-5.58	5.82	-3.84	17.13
	2010	29.54	-12.14	0.35	2.98	20.73		2010	26.36	-8.09	6.41	-5.86	18.82
	2013	30.87	-7.90	-1.60	1.38	22.75		2013	25.05	-9.87	8.89	-4.83	19.24
	2016	28.19	-5.42	-1.22	0.79	22.34		2016	29.70	-14.96	10.86	-3.94	21.66
20	1998	14.57	-11.51	-1.89	2.08	3.25	All	1998	7.56	-9.65	1.24	-0.21	-1.06
chemical	2001	13.03	-11.79	1.45	3.52	6.21		2001	9.33	-5.82	0.11	4.22	7.84
Products	2004	16.42	-10.41	3.71	2.62	12.34		2004	12.75	-5.75	0.20	3.97	11.17
	2007	12.86	-13.05	5.08	4.07	8.96		2007	12.93	-4.87	0.59	4.13	12.78
	2010	13.97	-18.06	8.52	9.34	13.77		2010	20.97	-8.27	-5.42	4.06	11.34
	2013	17.93	-15.91	7.41	6.59	16.02		2013	23.61	-7.91	-5.94	3.50	13.26
	2016	21.70	-9.71	2.47	1.51	15.97		2016	25.71	-6.50	-7.40	3.03	14.84

[a] All contributions represent growth rates in % relative to 1995.

[b] The total change in aggregate productivity is the sum of the contributions of survivors, entrants and exitors.

Table 6: Productivity Decomposition in the base year 1995 (t_1): survivors and exitors

Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_X^b	$S_X^{c,d}$	Ω_1^e	# Surv.	# Exits	Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_X^b	$S_X^{c,d}$	Ω_1^e	# Surv.	# Exits
10	1995	1998	0.11	86.63	0.51	13.37	0.16	19739	4327	24	1995	1998	-0.20	95.31	-0.04	4.69	-0.19	653	109
Food	1995	2001	0.11	72.20	0.28	27.80	0.16	14518	7850	Metals	1995	2001	-0.23	87.75	0.09	12.25	-0.19	543	201
	1995	2004	0.08	63.85	0.28	36.15	0.15	11586	9669		1995	2004	-0.22	81.33	-0.00	18.67	-0.18	451	274
	1995	2007	0.05	56.07	0.28	43.93	0.15	8999	11273		1995	2007	-0.22	78.28	-0.03	21.72	-0.18	390	327
	1995	2010	-0.01	49.89	0.32	50.11	0.16	6791	13317		1995	2010	-0.23	64.01	-0.09	35.99	-0.18	329	377
	1995	2013	-0.04	47.68	0.33	52.32	0.15	5537	14539		1995	2013	-0.13	57.78	-0.25	42.22	-0.18	299	402
	1995	2016	-0.06	44.26	0.32	55.74	0.15	4706	15214		1995	2016	-0.13	53.25	-0.23	46.75	-0.18	250	439
13	1995	1998	0.09	83.99	0.01	16.01	0.08	2205	475	28	1995	1998	-0.09	89.72	-0.24	10.28	-0.11	3264	894
Textile	1995	2001	0.10	74.57	0.03	25.43	0.08	1785	810	Machines	1995	2001	-0.08	85.13	-0.26	14.87	-0.10	2676	1401
	1995	2004	0.11	63.84	0.02	36.16	0.08	1483	1052		1995	2004	-0.05	76.74	-0.28	23.26	-0.11	2413	1624
	1995	2007	0.12	53.27	0.03	46.73	0.08	1162	1303		1995	2007	-0.06	71.29	-0.27	28.71	-0.12	2060	1929
	1995	2010	0.13	36.63	0.05	63.37	0.08	874	1578		1995	2010	-0.04	65.35	-0.25	34.65	-0.12	1552	2387
	1995	2013	0.12	30.37	0.05	69.63	0.07	705	1704		1995	2013	-0.14	55.43	-0.08	44.57	-0.11	1254	2657
	1995	2016	0.13	26.69	0.05	73.31	0.07	595	1792		1995	2016	-0.16	53.50	-0.06	46.50	-0.11	1069	2813
16	1995	1998	-0.21	87.77	-0.29	12.23	-0.22	3640	750	29	1995	1998	-0.60	98.15	0.06	1.85	-0.59	1158	281
Wood	1995	2001	-0.21	78.69	-0.26	21.31	-0.22	3002	1283	Automobiles	1995	2001	-0.20	32.80	-0.78	67.20	-0.59	974	433
	1995	2004	-0.21	67.71	-0.26	32.29	-0.23	2606	1620		1995	2004	-0.23	26.56	-0.73	73.44	-0.60	832	545
	1995	2007	-0.22	62.02	-0.24	37.98	-0.23	2251	1900		1995	2007	-0.21	26.19	-0.74	73.81	-0.60	717	633
	1995	2010	-0.20	54.13	-0.25	45.87	-0.23	1826	2298		1995	2010	-0.23	22.10	-0.71	77.90	-0.60	595	738
	1995	2013	-0.21	50.85	-0.25	49.15	-0.23	1540	2550		1995	2013	-0.22	21.46	-0.71	78.54	-0.60	519	806
	1995	2016	-0.19	44.51	-0.25	55.49	-0.22	1355	2695		1995	2016	-0.19	18.38	-0.70	81.62	-0.61	435	876
17	1995	1998	-0.15	92.76	-0.29	7.24	-0.16	1117	202	31	1995	1998	-0.20	85.78	-0.09	14.22	-0.18	3694	1252
Pulp/	1995	2001	-0.15	84.33	-0.19	15.67	-0.16	950	336	Furniture	1995	2001	-0.20	75.71	-0.11	24.29	-0.18	2974	1871
Paper	1995	2004	-0.14	74.31	-0.21	25.69	-0.16	817	449		1995	2004	-0.23	68.54	-0.09	31.46	-0.19	2473	2266
	1995	2007	-0.15	64.29	-0.17	35.71	-0.16	692	553		1995	2007	-0.22	62.11	-0.12	37.89	-0.18	2056	2600
	1995	2010	-0.13	55.79	-0.20	44.21	-0.16	577	646		1995	2010	-0.24	52.55	-0.12	47.45	-0.18	1269	3279
	1995	2013	-0.14	50.48	-0.17	49.52	-0.16	502	709		1995	2013	-0.23	43.94	-0.14	56.06	-0.18	956	3560
	1995	2016	-0.16	45.52	-0.17	54.48	-0.17	449	753		1995	2016	-0.22	37.16	-0.16	62.84	-0.18	730	3748
20	1995	1998	0.06	92.92	-0.24	7.08	0.03	1663	362	All	1995	1998	-0.13	90.34	-0.07	9.66	-0.13	4125	961
Chemical	1995	2001	0.07	83.20	-0.14	16.80	0.04	1362	623		1995	2001	-0.09	74.93	-0.15	25.07	-0.13	3198	1645
Products	1995	2004	0.07	69.47	-0.02	30.53	0.04	1172	778		1995	2004	-0.09	65.82	-0.14	34.18	-0.13	2648	2030
	1995	2007	0.08	62.98	-0.03	37.02	0.04	1007	930		1995	2007	-0.09	59.61	-0.14	40.39	-0.13	2148	2383
	1995	2010	0.14	54.91	-0.07	45.09	0.04	809	1095		1995	2010	-0.09	50.60	-0.15	49.40	-0.13	1624	2857
	1995	2013	0.11	49.58	-0.02	50.42	0.04	708	1181		1995	2013	-0.10	45.29	-0.14	54.71	-0.13	1335	3123
	1995	2016	0.00	44.20	-0.03	55.80	-0.01	646	1235		1995	2016	-0.11	40.83	-0.14	59.17	-0.14	1137	3285

[a] According to equation (20) this table reports measures of aggregate productivity and market shares for the group of surviving and exiting firms. The measures are always (and only) for the initial year 1995.

[b] The columns Ω_S and Ω_X denote the aggregate productivity of the firm groups survivors and exitors, respectively.

[c] The columns S_S and S_X denote the aggregated market shares of the firm groups survivors and exitors, respectively.

[d] S_S and S_X are given in %.

[e] Columns Ω_1 denotes the aggregate productivity for the initial year 1995.

Table 7: Productivity decomposition in t_2: survivors and entrants[a]

Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_E^b	$S_E^{c,d}$	Ω_2^e	# Surv.	# Entr.	Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_E^b	$S_E^{c,d}$	Ω_2^e	# Surv.	# no Entr.
10	1995	1998	0.07	91.11	0.73	8.89	0.13	19739	7962	24	1995	1998	-0.17	96.05	-0.01	3.95	-0.17	653	148
Food	1995	2001	0.01	77.13	0.24	22.87	0.07	14518	12879	Metals	1995	2001	-0.19	84.69	0.96	15.31	-0.01	543	227
	1995	2004	-0.04	70.85	0.27	29.15	0.05	11586	16691		1995	2004	-0.22	44.03	0.10	55.97	-0.04	451	374
	1995	2007	-0.02	68.02	0.25	31.98	0.06	8999	20054		1995	2007	-0.24	44.16	0.20	55.84	0.00	390	411
	1995	2010	-0.01	61.56	0.23	38.44	0.08	6791	21327		1995	2010	-0.22	37.98	-0.16	62.02	-0.18	329	441
	1995	2013	-0.09	55.98	-0.07	44.02	-0.08	5537	21396		1995	2013	-0.08	33.57	-0.14	66.43	-0.12	299	415
	1995	2016	-0.05	54.17	-0.16	45.83	-0.10	4706	18482		1995	2016	-0.07	25.18	-0.03	74.82	-0.04	250	333
13	1995	1998	0.15	86.75	0.19	13.25	0.15	2205	639	28	1995	1998	-0.12	88.71	-0.16	11.29	-0.13	3264	1283
Textile	1995	2001	0.19	80.68	0.16	19.32	0.19	1785	878	Machines	1995	2001	-0.13	83.54	0.09	16.46	-0.10	2676	1773
	1995	2004	0.22	75.60	0.18	24.40	0.21	1483	1011		1995	2004	-0.00	59.46	0.17	40.54	0.07	2413	3018
	1995	2007	0.24	69.52	0.22	30.48	0.24	1162	1062		1995	2007	0.04	55.20	0.08	44.80	0.06	2060	3322
	1995	2010	0.24	64.04	0.22	35.96	0.23	874	1002		1995	2010	0.06	51.23	0.00	48.77	0.03	1552	2738
	1995	2013	0.25	59.77	0.24	40.23	0.25	705	888		1995	2013	-0.02	51.40	0.15	48.60	0.06	1254	2279
	1995	2016	0.27	56.71	0.24	43.29	0.25	595	744		1995	2016	0.03	48.24	0.12	51.76	0.08	1069	2017
16	1995	1998	-0.20	89.70	-0.08	10.30	-0.19	3640	1178	29	1995	1998	-0.98	96.89	-0.06	3.11	-0.95	1158	397
Wood	1995	2001	-0.15	74.09	-0.26	25.91	-0.18	3002	1730	Automobiles	1995	2001	-0.15	36.66	-0.61	63.34	-0.44	974	605
	1995	2004	-0.06	71.34	0.00	28.66	-0.04	2606	2114		1995	2004	-0.25	22.40	-0.77	77.60	-0.65	832	714
	1995	2007	-0.09	66.72	-0.02	33.28	-0.06	2251	2460		1995	2007	-0.22	23.19	-0.78	76.81	-0.65	717	820
	1995	2010	-0.02	62.49	0.09	37.51	0.02	1826	2433		1995	2010	-0.24	30.05	-1.34	69.95	-1.01	595	770
	1995	2013	-0.01	61.96	0.12	38.04	0.04	1540	2260		1995	2013	-0.23	33.02	-1.40	66.98	-1.01	519	732
	1995	2016	0.04	58.74	0.17	41.26	0.09	1355	1873		1995	2016	-0.19	28.17	-1.39	71.83	-1.05	435	627
17	1995	1998	-0.11	91.64	-0.07	8.36	-0.11	1117	216	31	1995	1998	-0.15	86.97	-0.01	13.03	-0.14	3694	1678
Pulp/	1995	2001	-0.10	85.79	0.01	14.21	-0.09	950	327	Furniture	1995	2001	-0.18	79.35	0.03	20.65	-0.13	2974	2518
Paper	1995	2004	-0.08	78.89	-0.11	21.11	-0.09	817	405		1995	2004	-0.16	74.31	0.00	25.69	-0.12	2473	2922
	1995	2007	-0.08	73.43	-0.05	26.57	-0.07	692	443		1995	2007	-0.10	68.00	0.09	32.00	-0.04	2056	3251
	1995	2010	-0.06	65.61	-0.05	34.39	-0.05	577	421		1995	2010	-0.11	62.77	0.06	37.23	-0.04	1269	2120
	1995	2013	-0.01	61.26	-0.05	38.74	-0.02	502	411		1995	2013	-0.15	60.43	0.07	39.57	-0.06	956	1756
	1995	2016	-0.02	54.64	-0.04	45.36	-0.03	449	391		1995	2016	-0.21	61.28	0.07	38.72	-0.10	730	1382
20	1995	1998	0.09	89.44	-0.09	10.56	0.07	1663	487	All	1995	1998	-0.16	90.81	0.05	9.19	-0.15	4125	1554
Chemical	1995	2001	0.07	78.06	0.13	21.94	0.08	1362	683		1995	2001	-0.07	75.55	0.08	24.45	-0.07	3198	2402
Products	1995	2004	0.07	70.65	0.20	29.35	0.11	1172	864		1995	2004	-0.06	63.06	0.00	36.94	-0.06	2648	3123
	1995	2007	-0.01	64.55	0.13	35.45	0.04	1007	900		1995	2007	-0.05	59.20	0.01	40.80	-0.05	2148	3635
	1995	2010	-0.07	61.85	0.16	38.15	0.02	809	928		1995	2010	-0.05	55.29	-0.09	44.71	-0.10	1624	3575
	1995	2013	-0.05	48.82	0.09	51.18	0.02	708	902		1995	2013	-0.04	51.80	-0.11	48.20	-0.10	1335	3448
	1995	2016	-0.01	50.77	0.04	49.23	0.01	646	832		1995	2016	-0.02	48.66	-0.11	51.34	-0.10	1137	2964

[a] According to equation (21) this table reports measures of aggregate productivity and market shares for the group of surviving and entering firms. The values always represent measures concerning the second time period t_2.

[b] The columns Ω_S and Ω_E denote the aggregate productivity of the firm groups survivors and entrants, respectively.

[c] The columns S_S and S_E denote the aggregated market shares of the firm groups survivors and entrants, respectively.

[d] S_S and S_E are given in %.

[e] Column Ω_2 denotes the aggregate productivity for the second time period t_2.

5.4 Productivity differences

I now present the results of the analysis of productivity differences among different firm groups. As described in Section 4.4, I proceed first by comparing the empirical cumulative density function (ECDF) between the different groups. In a second step I discuss the results of the Kolmogorov-Smirnov (KS) test, measuring the statistical significance in the difference between the distributions of two compared groups. Note that for the comparison of the productivity distribution of the firm groups I always compare firms for a chosen entry and exit cohort. To define firm survival, entry and exit for a specific cohort t I rely on the definition in Section 3.2. That is, other than in the previous section, where I defined entry and exit for longer time spans, I here consider firms in a specific year that where identified in a yearly perspective as survivors, entrant or exitors. In Section 5.4.1 I discuss productivity differences between surviving and entering firms; in Section 5.4.2 between surviving and exiting firms; in Section 5.4.3 between entering and exiting firms; in Section 5.4.4 between surviving and failing entrants; and in Section 5.5 I discuss the relation between productivity and firm size and compare different firm size groups.

5.4.1 Survivors vs. entrants

In order to compare the productivity distributions of survivors and entrants I consider the entry cohorts 1995, 2000, 2005, 2010 and 2015. Figure 8 provides the ECDF's of both groups for the respective cohort. The dashed (solid) line represents the ECDF of the group of survivors (entrants). It can be seen that for all cohorts, the ECDF of the entrants is located slightly to the left compared to the EDCF of the survivors. For the entry cohort 2015, the gap between both CDF's is larger, especially in the area where the slope of both curves rapidly increases, i.e. in the area of the highest density of observations. Generally, the graphs indicate that the productivity distribution of survivors stochastically dominates the productivity distribution of entering firms. This is the case since for all percentiles (y-axis) the productivity level (x-axis) of survivors is at least as high as the productivity level of the entering firms. Table 8 shows the results of the KS-test on both distributions. It can be seen that the two-sided test rejects at a high level of significance for all cohorts the null hypothesis of equality of both distributions. Moreover, the hypothesis that surviving firms have higher productivity distribution compared to the group of entering firms can not be rejected. In other words, the ECDF of survivors is right to the one of the ECDF of entrants and, thus, the productivity distribution of surviving firms stochastically dominates the distribution of entering firms. The last column of Table 8 reports the difference in the median value of productivity between both groups at the median. The figures show that the median level of productivity is higher for the group of survivors compared to the group of entrants, with a difference ranging between 7.0% in 2000 to 17.8% in 2010. Note that these results of productivity differences between surviving and entering firms confirm those presented in Fariñas and Ruano (2005) for Spain and Wagner (2010) for Germany.

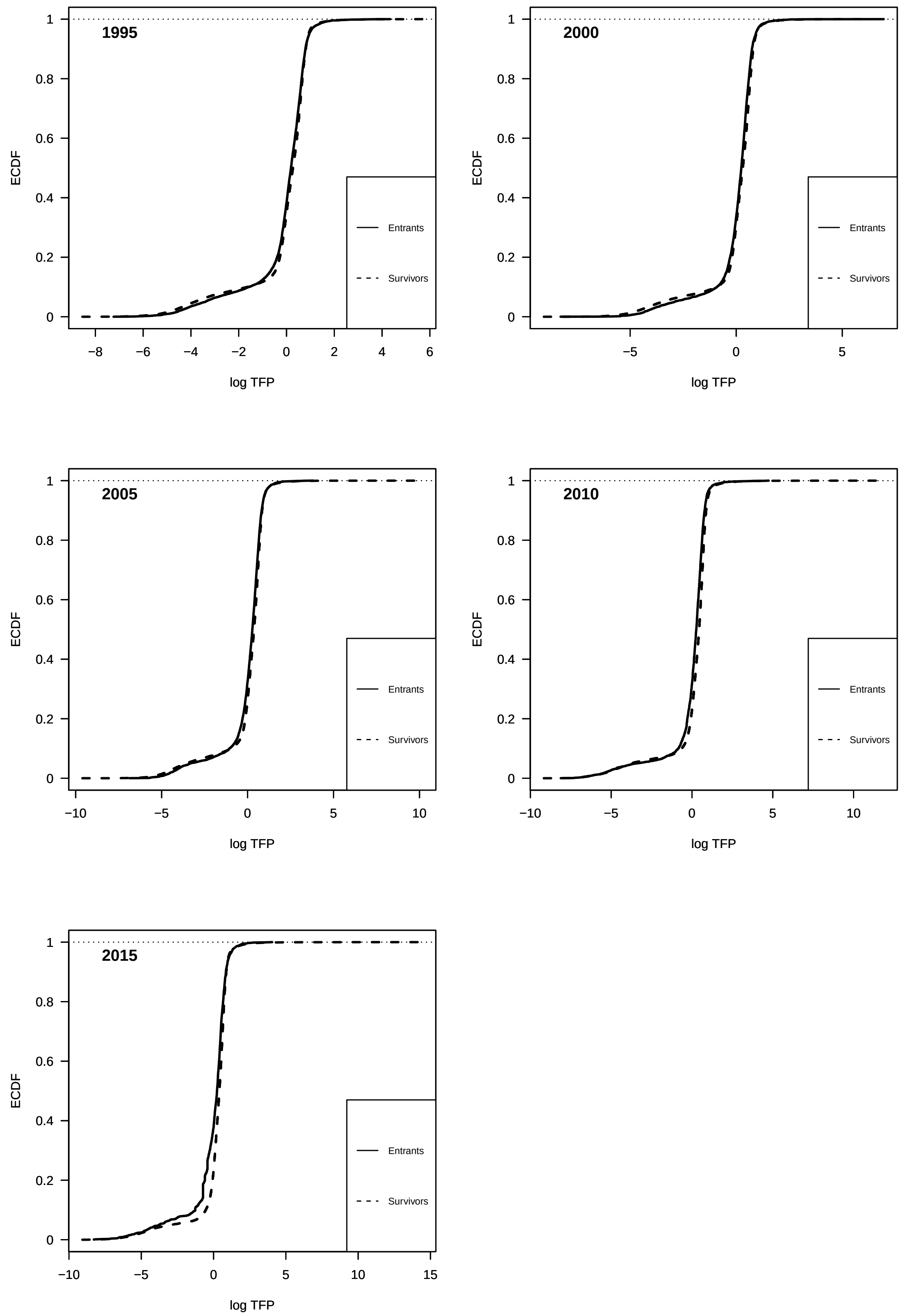

Figure 8: The distribution of productivity for survivors vs. entrants

Table 8: Productivity differences between survivors and entrants

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to survivors		Median difference
Year	# of Entrants	# of Survivors	Statistic	p-value	Statistic	p-value	med_S-med_E
1995	6346	41341	0.058	0.000	0.014	0.129	0.082
2000	5534	42693	0.090	0.000	0.013	0.175	0.070
2005	4755	47140	0.080	0.000	0.009	0.466	0.099
2010	4547	42200	0.150	0.000	0.007	0.633	0.178
2015	1515	30317	0.166	0.000	0.007	0.853	0.165

5.4.2 Survivors vs. exitors

Next I compare productivity differences between the group of survivors and exitors. Figure 9 presents the productivity ECDF's of both groups for the considered cohorts. Here, the dashed (solid) line represents the ECDF of survivors (entrants). Similar to the previous case, for all years, the productivity ECDF of the group of survivors is located slightly to the right. This indicates that for each percentile the level of productivity of survivors is higher compared to those of exitors. Only for the year 2010, exitors and survivors do not seem to exhibit a significant difference in productivity. Generally, however, the graphs indicate that the productivity distribution of surviving firms stochastically dominates the productivity distribution of exiting firms. Table 9 provides the results from the corresponding KS-test. It can be seen that for all years the two-sided tests rejects the null hypothesis of equality of both distributions, at a high level of significance. Considering the one-sided test, the hypothesis that the difference is favorable to surviving firms can not be rejected except for the year 2010, which goes in line with the graphical finding. I conclude that for all considered cohorts, except 2010, the productivity distribution of survivors stochastically dominates the productivity distribution of exitors. The median difference, sustains this difference where survivors for all years have a higher level of the median productivity level, with a difference ranging between 4.3% in 2015 and 12.4% in 2015. For the year 2010, the median difference in productivity is slightly favorable for exitors, given by 2.5%. The results here again go in line with those presented in Fariñas and Ruano (2005) and Wagner (2010).

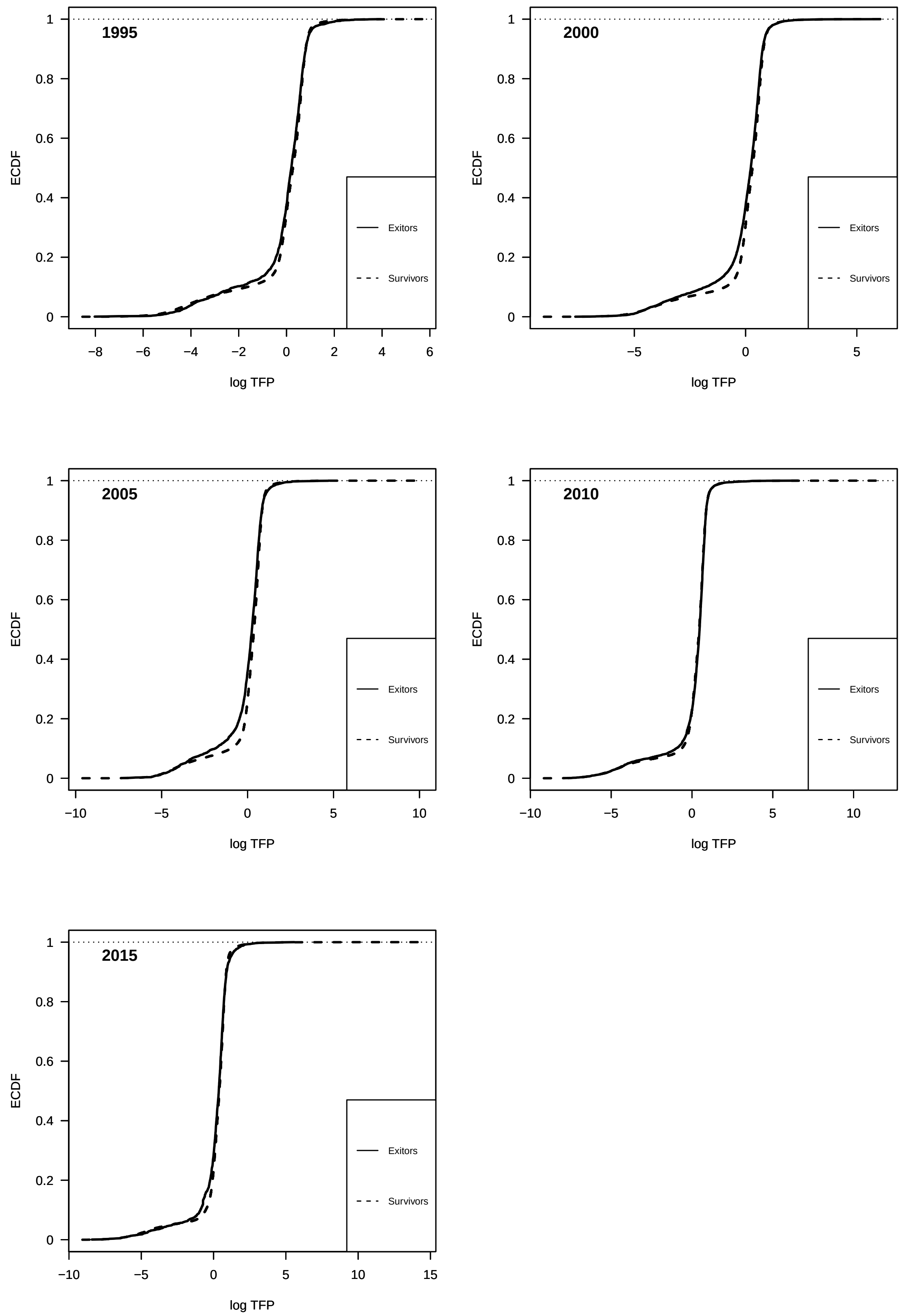

Figure 9: The distribution of productivity for survivors vs. exitors

Table 9: Productivity differences between survivors and exitors

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to survivors		Median difference
Year	# of Exitors	# of Survivors	Statistic	p-value	Statistic	p-value	med_S-med_X
1995	2268	41341	0.052	0.000	0.011	0.613	0.068
2000	5389	42693	0.083	0.000	0.003	0.900	0.078
2005	2492	47140	0.101	0.000	0.010	0.631	0.124
2010	4998	42200	0.026	0.004	0.026	0.002	-0.025
2015	1608	30317	0.062	0.000	0.023	0.192	0.043

5.4.3 Entrants vs. exitors

Figure 10 illustrates the graphical comparison between the productivity ECDF's of both entrants and exitors, where the dashed (solid) line represents the ECDF of exitors (entrants). For the entry/exit cohorts 1995, 2000, and 2005 exitors seem to have lower productivity levels especially for low levels of productivity, where the dashed line lies slightly above the solid line. For higher productivity level both CDFs collapse. Contrarily, for the year 2015 the CDFs of exitors lies relatively constantly below the CDF of entrants, which indicates stochastic dominance in favor of the productivity distribution of exitors. For the year 2010 the curves intersect, that is, in the region of lower productivity, the CDF of entrants lies below the CDF of exitors, while for higher productivity levels the opposite is the case.

The corresponding KS-test results are reported in Table 10. The two-sided test reveals that except for the year 1995 the null hypothesis of equality of both distribution can be rejected. The one-sided test for the hypothesis that the productivity difference is favorable to exitors can not be rejected for the cohorts 1995, 2010 and 2015. Except for the year 1995, this implies that the productivity distribution of the group of exitors for these cohorts stochastically dominates the distribution of entrants. Note that the test result for the year 1995, is ambiguous since on the one hand equality of both distribution can not be rejected, and on the other hand, the one-sided test indicates a difference in productivity favorable for exitors. This might steam from the fact that the distributional difference for this year is very small as also indicated by the median difference in the last column. Instead, the hypothesis of can be rejected for the cohorts 2000 and 2005. This supports the graphical analysis, where especially for the years 2000 and 2005 for some regions a slightly favorable difference for the group of entrants can be seen. The median difference reflects test results: For the years 1995, 2010 and 2015 the median productivity level of exitors is slightly higher compared to exitors, with a difference ranging between -1.4% in 1995 to -20.3% in 2010. Fariñas and Ruano (2005) also find very small differences between entering and exiting firms, however, their general pattern indicates a higher productivity levels for entering firms.

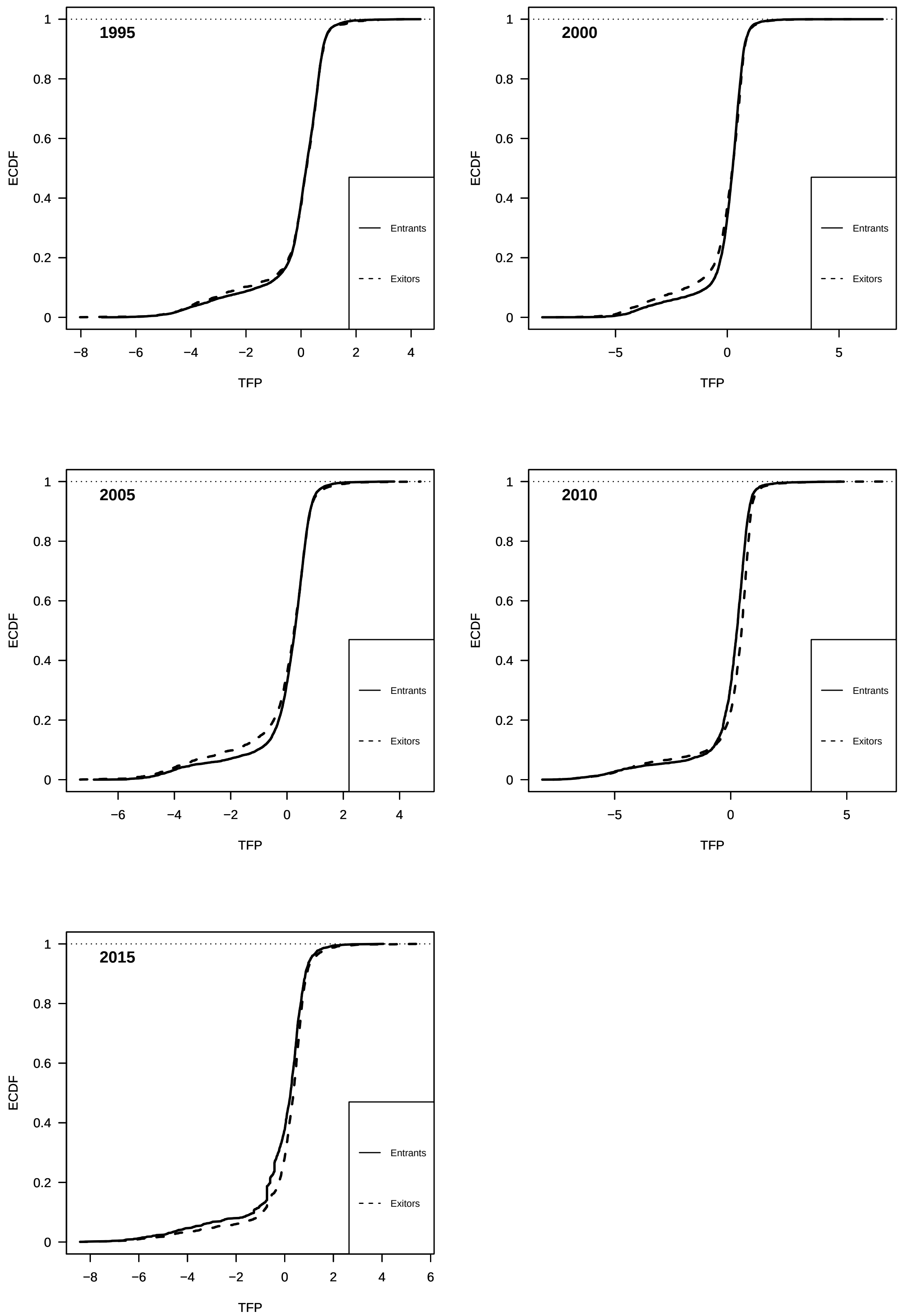

Figure 10: The distribution of productivity for entrants vs. Exitors

Table 10: Productivity differences between entrants and exitors

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to exitors		Median difference
Year	# of Entrants	# of Exitors	Statistic	p-value	Statistic	p-value	med_E-med_X
1995	6346	2268	0.017	0.704	0.017	0.371	-0.014
2000	5534	5389	0.051	0.000	0.051	0.000	0.008
2005	4755	2492	0.047	0.002	0.047	0.001	0.025
2010	4547	4998	0.171	0.000	0.014	0.419	-0.203
2015	1515	1608	0.111	0.000	0.001	0.999	-0.123

5.4.4 Surviving entrants vs. failing entrants

I now compare surviving vs. failing entrants. This is an interesting comparison since in the literature firms' initial productivity level is often seen as a determinant factor for firms' future success (Hopenhayn, 1992). I define surviving entrants (SE) as firms entering at t and being registered as survivors in $t+1$. According to the definition of entry and survival, this means that surviving entrants are active for at least three consecutive years. Instead, failing entrants (FE) are defined as firms entering at t and being identified as exitor at $t+1$. Note that a firm identified as exitor at t is still active in t but effectively exits in $t+1$. By this definition, failing entrants are only active for two consecutive years. I analyse the entry cohorts for the years 1995, 2000, 2005, 2010 and 2014.[21] Figure 11 provides the graphical comparison of the productivity ECDF's of both groups. Here the dashed (solid) line represents the ECDF of surviving (failing) entrants. The graphs show, especially for the years 1995, 2000, 2005, and 2010 a relatively consistent picture, where the productivity ECDF of surviving entrants is located to the right relative to the ECDF of failing entrants. For these years the productivity levels of surviving entrants are higher compared to those of failing entrants, indicating that the distribution of surviving entrants stochastically dominates the distribution of failing entrants. It seems, however, that the productivity differences between both groups become smaller over time since the ECDFs move closer together. For the exit cohort 2014 I observe an intersection between both ECDFs, where for very low levels of productivity failing entrants exhibit a higher productivity (the ECDF located to the right). After the intersection, with the steep increase of the ECDFs, the productivity level of failing entrants lies below that of surviving entrants (located to the left). In this case, for none of the distributions stochastic dominance can be assessed. Table 11 provides the corresponding KS-test results. The null hypothesis of the two-sided test, i.e. equality between both distributions, is rejected for all entry cohorts, except 2014, at the 5% significance level. The one-sided test, testing for the hypothesis of a difference favorable for failing entrants is rejected at the 1% significance level for the years 1995 - 2010, but cannot be rejected at a reasonable significance level for the year 2014. This means that for the entry cohort 1995, 2000 and 2010, the productivity distribution of surviving entrants stochastically dominates the distribution of failing entrants. As before, these findings support the results with respect to productivity differences between surviving and failing entrants presented in Fariñas and Ruano (2005) for Spain and in Wagner (2010) for Germany. The last column in Table 11 shows productivity differences at the median of both distributions. It can be seen that median productivity level of surviving entrants is higher, and the difference ranges between 19.2% in 1995 and 5.7% in 2014.

[21] Note that I chose 2014 instead of 2015 as we did in the previous cases, since surviving entrants need to be defined as survivors in $t+1$. Survivors, themselves are defined as being active in $t-1$, t and $t+1$. However, since 2016 is the last year of my sample period, the last year for which survivors can be identified is the year 2015.

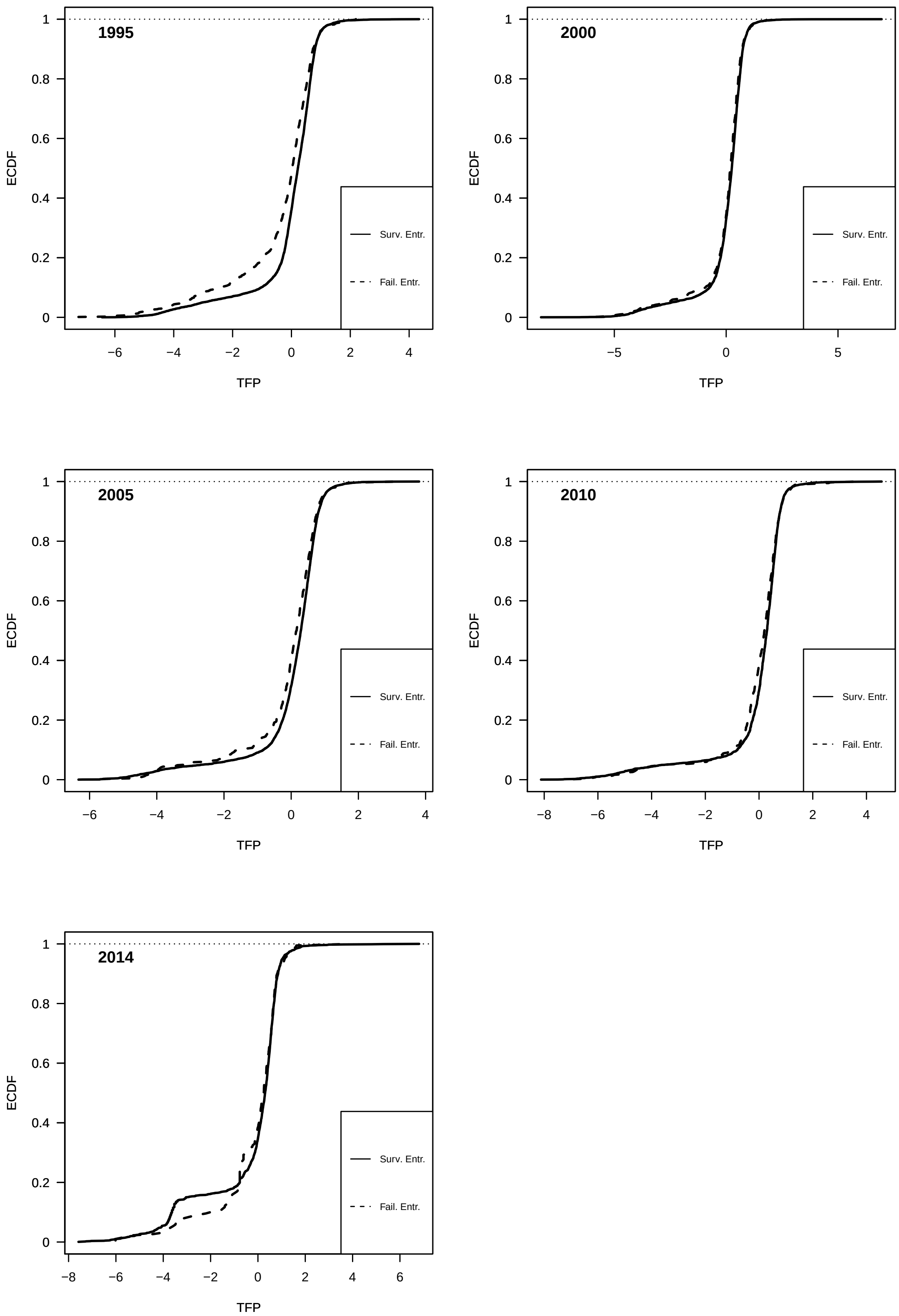

Figure 11: The distribution of productivity for surviving entrants vs. failing entrants

Table 11: Productivity differences between surviving and failing entrants

Year	Observations: # of Surviving Entrants	Observations: # of Failing Entrants	Two-sided test: Equality of distributions: Statistic	Two-sided test: Equality of distributions: p-value	One-sided test: Difference favorable to failing entrants: Statistic	One-sided test: Difference favorable to failing entrants: p-value	Median difference: med_{SE}-med_{FE}
1995	5244	804	0.143	0.000	0.143	0.000	0.192
2000	4697	618	0.063	0.025	0.063	0.013	0.063
2005	4125	412	0.099	0.001	0.099	0.001	0.126
2010	3658	619	0.089	0.000	0.089	0.000	0.095
2014	1665	188	0.075	0.293	0.071	0.186	0.057

5.5 Productivity differences with respect to firm size

To investigate the relation between firms' productivity and size I first take a look at a correlation matrix reporting correlations of log productivity, $\hat{\omega}_{nt}$, and the log input demand variables, x^k_{nt}, x^l_{nt}, and x^m_{nt}, as well as the binary firm status variables entry and exit, e^+_{nt} and e^-_{nt}. In a second step I then analyze - similar to the previous comparisons - the productivity distributions of different firm size groups.

Table 12 presents the correlations between firms' log productivity and the log of capital, labor and materials demand, as well as the correlation between firms' productivity and the dummy variables entry and exit. All correlations are reported based on the whole industry and for each 2-digit industry separately. Surprisingly, the correlations I find are all very small. Interpreting only the sign, firms' productivity and the use of capital is positively related when the correlation is calculated based on the whole sample, while for industries "wood", "pulp/paper", and "furniture" a slightly negative correlation is found. Regarding the correlation between productivity and labor I find a negative overall correlation but a slight positive correlation for the industries "food", "textiles", "chemical products", "metals" and "automobiles". The figures also show a negative overall correlation (as well as for all 2-digit industries) between productivity and firm entry, and productivity and form exit, respectively. Generally the figures indicate that there is no strong relation between firms' size (measured by the different input variables) and and their (log) productivity.

Table 12: Correlations

Industry	$cor(\hat{\omega}_{nt}, x^k_{nt})$	$cor(\hat{\omega}_{nt}, x^l_{nt})$	$cor(\hat{\omega}_{nt}, x^m_{nt})$	$cor(\hat{\omega}_{nt}, e^+_{nt})$	$cor(\hat{\omega}_{nt}, e^-_{nt})$	Obs.
All	0.01	-0.01	-0.07	-0.03	-0.01	1118540
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
10 food	0.03	0.02	-0.01	-0.03	-0.01	643411
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
13 textiles	0.06	0.03	-0.08	-0.04	-0.04	52123
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
16 wood	-0.01	-0.06	-0.17	-0.03	0.00	100259
	(0.00)	(0.00)	(0.00)	(0.00)	(0.95)	
17 pulp/paper	-0.02	-0.03	-0.15	-0.02	-0.02	25724
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	
20 chem. products	0.03	0.01	-0.10	-0.03	-0.04	43139
	(0.00)	(0.09)	(0.00)	(0.00)	(0.00)	
24 metals	0.03	0.04	-0.09	-0.04	-0.05	17201
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
28 machines	0.00	-0.01	-0.21	0.00	-0.02	101411
	(0.17)	(0.03)	(0.00)	(0.53)	(0.00)	
29 automobiles	0.09	0.06	-0.09	-0.05	-0.04	32740
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
31 furniture	-0.01	-0.02	-0.19	-0.03	-0.02	102532
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	

Note: p-values are given in parenthesis.

Analogously to the investigation of productivity differences between survivors, entrants and exitors, I conduct a comparison of the ECDFs of different firm size groups, graphically and by the application of the KS-test. For this purpose I cluster firms into three groups: 1) small firms, with a number of employees between 1 and 50; 2) medium firms, with a number of employees between 51 and 250; and 3) big firms with a number of employees of more than 250. Figure 12 reports the ECDF of small, medium and big firms, represented by the solid, dashed and dotted lines, respectively. As before, I compare size differences for the years 1995, 2000, 2005, 2010 and 2015. Considering the graphs it seems that small firms are less productive for lower percentiles, since the ECDF of small firms is located to the left compared to medium and big firms. However, at around the 20th percentile the ECDF of small firms intersects the ECDF of both medium and big firms and is then consistently located to the right, compared to the ECDFs of medium and big firms, indicating for higher productivity levels within this range. Also it seems that the ECDF of medium firms is slightly located to the right for higher percentiles compared to the ECDF of big firms. This indicates that the productivity level of big firms is lower compared to smaller firms at higher ranges of the productivity distribution. However, it can also be seen that for very low productivity levels, the ECDFs of medium and big firms intersect, which does not allow to conclude for stochastical dominance of the distributions.

The results of application of the KS-test for the comparison between the productivity distributions between small and medium, small and big, as well as medium and big firms are presented in Table 13, 14 and 15, respectively. For all comparisons and years the null hypothesis of equality of the considered distributions is highly rejected. Furthermore, the one-sided test, testing the hypothesis that the difference is favorable to the respective larger firm group is also strongly rejected for each compared group (and year). From these tests I conclude that none of the productivity distributions of the compared size groups stochastically dominates its counterpart. Analysing firm-level data of African manufacturing, industries Van Biesebroeck (2005) also finds no stochastical dominance of the (total factor) productivity distribution of large firms with respect to small firms. Instead, analysing only labor productivity he shows that the labor productivity distribution of large firms stochastically dominate the one smaller firms. For the case of US and Canadian firms Leung et al. (2008) find a positive relation between firm size and labor and total factor productivity in manufacturing and non-manufacturing industries.

Table 13: Productivity differences between small and medium

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to medium		Median difference
Year	# of small firms	# of medium firms	Statistic	p-value	Statistic	p-value	med_{medium}-med_{small}
1995	46652	3098	0.267	0.000	0.267	0.000	-0.374
2000	51209	3133	0.260	0.000	0.260	0.000	-0.340
2005	50848	3046	0.292	0.000	0.292	0.000	-0.357
2010	49198	2847	0.297	0.000	0.297	0.000	-0.385
2015	30330	2706	0.303	0.000	0.303	0.000	-0.359

Table 14: Productivity differences between small and big

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to big firms		Median difference
Year	# of small firms	# of big firms	Statistic	p-value	Statistic	p-value	med_{big}-med_{small}
1995	46652	909	0.434	0.000	0.434	0.000	-0.511
2000	51209	941	0.407	0.000	0.407	0.000	-0.480
2005	50848	986	0.482	0.000	0.482	0.000	-0.531
2010	49198	843	0.519	0.000	0.519	0.000	-0.594
2015	30330	813	0.504	0.000	0.504	0.000	-0.596

Table 15: Productivity differences between medium and big

	Observations		Two-sided test: Equality of distributions		One-sided test: Difference favorable to big firms		Median difference
Year	# of medium firms	# of big firms	Statistic	p-value	Statistic	p-value	med_{big}-med_{medium}
1995	3098	909	0.173	0.000	0.173	0.000	-0.137
2000	3133	941	0.189	0.000	0.189	0.000	-0.14
2005	3046	986	0.215	0.000	0.215	0.000	-0.174
2010	2847	843	0.257	0.000	0.257	0.000	-0.208
2015	2706	813	0.309	0.000	0.309	0.000	-0.237

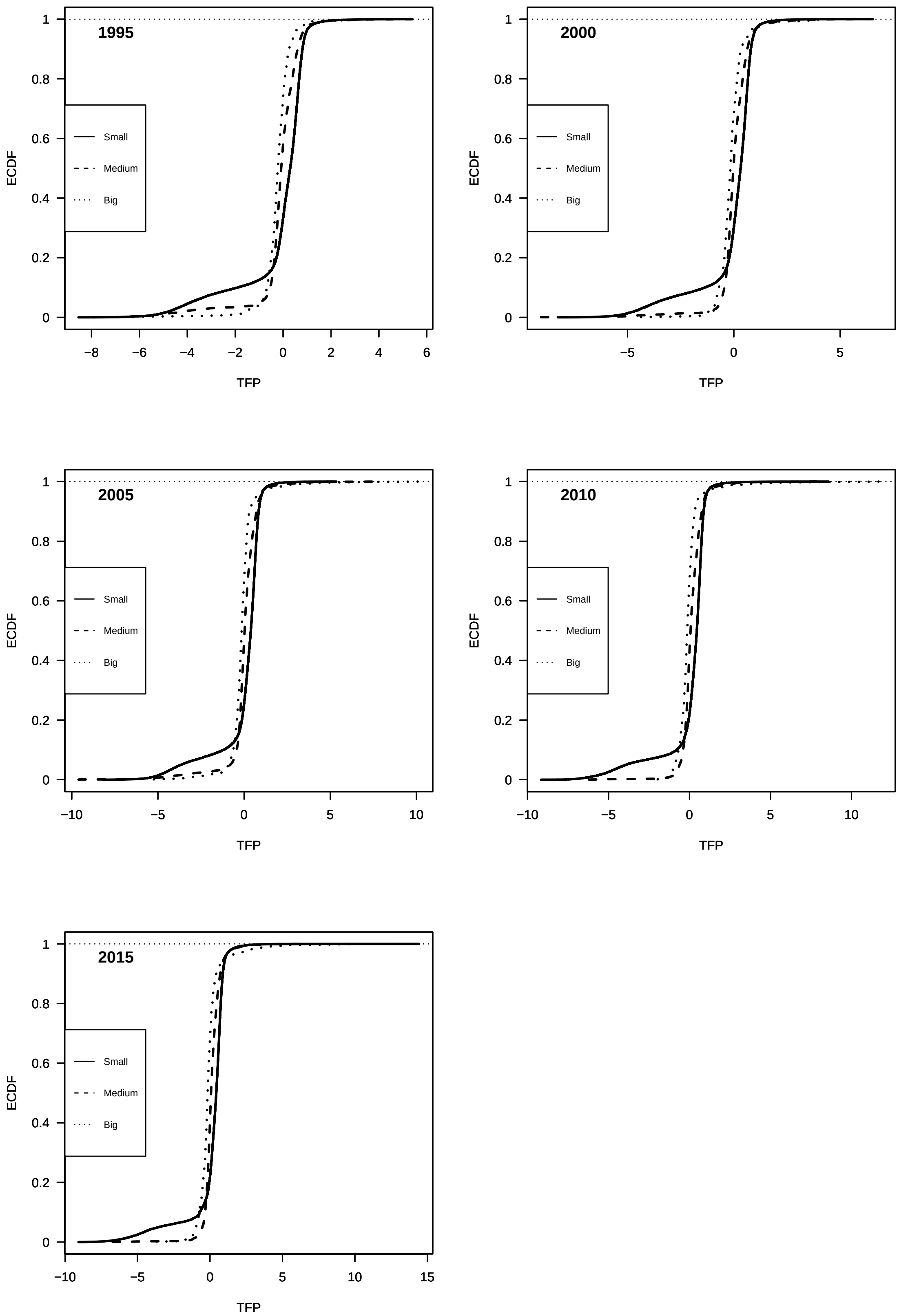

Figure 12: The distribution of productivity for small vs. medium vs. big firms

6 Limits

My analysis is limited in various ways. To begin with limitations related to the estimation of the translog production function and, hence, related to the measurement of productivity. First, the production function I use to estimate firm-level productivity implies the same production technology for all firms, industries and years up to some additive random terms. This is certainly a limitation given the different industries considered and the long sample period during which, in my setting, the parameters are supposed to either fixed or random, so that heterogeneities can be subsumed in the additive random term. Technological differences are studied in more details by De Monte and Koebel (2020) using the same data. Second, I assume Hicks neutral technical change, which implies that the productivity term ω_{nt} is additive in the expression of the production function. Third, since I use deflated firm sales as gross output production, my productivity measure does probably not fully reflect firms' efficiency because of price effects that occur within the considered 2-digit industries, if firms practice product differentiation, and/or if firms detain market power (Van Biesebroeck, 2008). Fourth, to estimate the production function I only control for a potential simultaneity bias. However, as Olley and Pakes (1996) argues, there might also be a selection bias due to endogenous firm exit, which is neglected in the adopted estimation procedure.

Moreover, my analysis is limited with respect to the identification of firm entry and exit. My data set does not allow to definitely identify firm entry and exit in a legal sense. That is, firms may be inactive for a certain period and then reactivate their business. I count such cases as entry (reactivation of activity) and exit (temporal inactivity). Also, I am not able to identify mergers and acquisitions. That is, if a firm acquires a second firm I consider this as an exit of the acquired firm - even if it continues to produce.

7 Conclusion

This paper investigates French firm-level data for different manufacturing industries, covering the period 1994 - 2016. My main interests are (i) the evolution of aggregate productivity related to firm entry and exit, (ii) productivity persistence and (iii) productivity differences between different firm groups such as survivors, entrants and exitors as well as between firm size groups. My results show that aggregate productivity has increased over time for most of the investigated 2-digit industries. Furthermore, aggregate productivity change is mainly driven by productivity improvements of surviving firms (within change). See Baily et al. (1992), Griliches and Regev (1995), Pavcnik (2002) and Melitz and Polanec (2015) who also find an important within contribution of surviving firms to aggregate productivity growth. Furthermore, I find that the contribution of aggregate productivity changes via market shares reallocation of surviving firms (between change) is for most industries negative, indicating inefficient allocation of production (Haltiwanger, 2011). Moreover, according to my measures the contribution of firm entry and exit to aggregate productivity growth varies substantially among industries. In comparison to the contribution of surviving firms, entry and exit plays especially for the industries "food", "metals" and "automobiles" an important role with respect to aggregate productivity growth. While in the metals industry entrants contribute substantially positively to aggregate productivity growth, for the automobiles industry I measure a strong negative contribution of the same group. Similarly, while I measure for the food industry a negative contribution of exitors with respect to the industries total aggregate productivity change, the group of exitors in the automobile industry has contributed considerably positively. Generally, my results suggest that the group of entering and exiting firms, however, contribute less to aggregate productivity growth compared to the contribution of surviving firms, which confirms findings in Baily et al. (1992) and Melitz and Polanec (2015). For industries that experienced a negative contribution of exitors this implies that relatively high productive firms have exited the market and, hence, exiting firms contribute negatively to the aggregate productivity growth. Haltiwanger (2011) defines a well-functioning economy as an economy revealing allocative efficiency, i.e. more productive firms tend to hold more market shares and market shares wander from less to high productive firms over time. In this sense, those industries that exhibit a negative contribution by the group of exitors indicate to be badly-functioning economies, since in these cases the group of exiting firms has a higher aggregate productivity compared to the group of surviving firms, where the latter group takes over the largest part of the left market shares, implying that market shares wander by trend from higher to low productive firms.

In order to investigate productivity differences with respect to different firm groups I apply the concept of stochastic dominance, following Fariñas and Ruano (2005). My results show that over all industries survivors have higher productivity levels compared to entering and exiting firms. Entrants have higher productivity compared to exitors. And surviving entrants have higher productivity compared to failing entrants. These results go in line with Fariñas and Ruano (2005) analysing Spanish manufacturing industries and a similar study by Wagner (2010) for the case of German manufacturing industries. Furthermore, these findings confirm model assumptions concerning firm selection, i.e. the process of firm entry and exit, in Hopenhayn (1992). My analysis of productivity differences with respect to different firm size groups (small, medium and large firms), shows that within a range of lower productivity, large firms have higher productivity compared to small and medium firms. However, within a range of higher productivity levels, small firms turn out to be more productive with respect to medium and large firms. These findings go in line with Van Biesebroeck (2005) for the case of firms active in African manufacturing industries but are contradicting to results shown in Leung et al. (2008), where US and Canadian manufacturing and non-manufacturing firms are analysed.

Appendix A: Data

A1: Merging of the data sets FICUS and FARE

For my analysis I merge the two fiscal firm-level data sets FICUS and FARE, covering the periods from 1994 to 2007, and 2008 to 2016, respectively. Both in FICUS and FARE firms are classified by a 4-digit industry nomenclature "NAF" (nomenclature d'activité française). However, from FICUS to FARE this industry nomenclature has significantly changed. In FICUS, the nomenclature was organized according to "NAF 1", while in FARE the nomenclature is organized according to "NAF 2". In this study I treat one single data set, 1994 - 2016, by establishing consistency in the industry nomenclature NAF 2 throughout the whole period. That is, I assign the current 4-digit industry nomenclature NAF 2 retrospectively for all firm observations from FICUS. For firms that are observed both in FICUS and FARE or only in FARE the 4-digit industry according to NAF 2 they belong to is known. However, for firms that have exited the market before 2008 I do not know to which NAF 2 4-digit industry they would have belonged to if they had continued their activity. To also classify these firms by the NAF 2 4-digit nomenclature I use the following methodology. I first only look at firms that are observed in both data sets FICUS and FARE. From these observations I build a transition matrix where each row represents a 4-digit industry according to NAF 1 and each column represents a 4-digit industry according to NAF 2. Each cell of the transition matrix contains the number of firms transiting from a specific 4-digit industry in FICUS (NAF 1) to the new 4-digit industry in FARE (Naf 2). Table 16 shows an exemplifying transition matrix, where I chose the NAF 1 4-digit industries 201A - 205C, i.e. the wood working industries. For instance it can be seen that there are 2060 firms observed that were classified in FICUS in 201A (first row) and in FARE in the industry 1610 (third columns), while there are only 46 observations that were classified in 201A and in FICUS in 0220 (first column). From these observed transition frequencies I then calculate the transition probabilities by simply dividing each element of the matrix by the sum of its corresponding row. That is, the NAF 1 - NAF 2 transition probabilities are calculated by

$$p_{ij} = \frac{\sum_{n \in i,j}^{N_i} \mathbf{1}_{[n \in i \text{ and } n \in j]}}{\sum_{n=1 \in i}^{N_i} \mathbf{1}_{[n \in i]}}, \tag{27}$$

where n is a firm observed in both FICUS and FARE, i and j are specific 4-digit industries according to NAF 1 and NAF 2, respectively. $\mathbf{1}$ is an index variable equal to 1 if the condition in parenthesis is fulfilled. Table 17 contains the transition probabilities according to the observed transitions Table 16. It can be seen that those 4-digit transitions between FICUS and FARE that were more frequently observed obtain accordingly higher probabilities. In a second step, firms only observed in FICUS belonging to a specific NAF 1 4-digit industry, are assigned to a NAF 2 4-digit, by drawing from a discrete probability distribution which corresponds to the row in the probability transition matrix, i.e. the NAF 1 4-digit industry a firm belongs to and its potential transition possibilities.

Table 16: FICUS - FARE: Observed transition frequencies

	NAF 2																		
NAF 1	0220	1392	1610	1621	1622	1623	1624	1629	2223	2512	3101	3109	3319	4329	4332	4391	4399	5610	9524	Total
201A	46	0	2060	5	6	22	35	12	0	0	0	7	0	0	25	24	9	5	0	2256
201B	0	0	498	0	0	0	0	0	0	0	0	0	0	17	4	36	24	0	0	579
202Z	0	0	0	108	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	112
203Z	0	7	33	0	15	1880	8	8	41	26	0	41	0	6	1005	386	34	0	0	3490
204Z	0	0	17	0	0	4	857	6	0	0	0	0	35	0	6	0	0	0	0	925
205A	4	16	10	4	0	21	5	1215	0	0	12	317	0	0	87	0	4	10	156	1861
205C	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	86

Table 17: FICUS - FARE: Transitions probabilities

	NAF 2																			
NAF 1	0220	1392	1610	1621	1622	1623	1624	1629	2223	2512	3101	3109	3319	4329	4332	4391	4399	5610	9524	Total
201A	0.02	0.00	0.91	0.00	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	1.00
201B	0.00	0.00	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.06	0.04	0.00	0.00	1.00
202Z	0.00	0.00	0.00	0.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	1.00
203Z	0.00	0.00	0.01	0.00	0.00	0.54	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.29	0.11	0.01	0.00	0.00	1.00
204Z	0.00	0.00	0.02	0.00	0.00	0.00	0.93	0.01	0.00	0.00	0.00	0.00	0.04	0.00	0.01	0.00	0.00	0.00	0.00	1.00
205A	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.65	0.00	0.00	0.01	0.17	0.00	0.00	0.05	0.00	0.00	0.01	0.08	1.00
205C	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

A2: Supplementary information on 2-digits industries and observations

Table A1 below provides the list of considered 2-digit industries with the corresponding number of firms as well as the total number of observations.

Table 18: Description 2-Digit Manufacturing Industries

Industry	Description	# Firms	# Obs.
10	Manufacture of food processing	101,766	702,388
13	Manufacture of textiles	6,821	56,685
16	Manufacture of wood, products of wood and cork	13,199	108,751
17	Manufacturing of pulp paper and paperboard	2,807	28,055
20	Manufacturing of chemical products	5,173	46,945
24	Manufacturing of basic metals	2,056	18,683
28	Manufacturing of machines and equipment	13,137	109,306
29	Manufacturing of automobiles	4,010	35,597
31	Manufacture of furniture	15,545	110,451
Total		164,514	1,216,861

A3: Missing values treatment

First of all, is noteworthy to mention again that I only keep firms reporting at least one employee. In doing so, I also drop all missing values for the variable log labor (x^l_{nt}). For the estimation of firm-level productivity in the framework of a translog production function it is only possible to obtain productivity estimates of those firms reporting positive values for the variables gross output production as well as the input factors capital, labor and materials. Since I relate firm productivity to entry and exit I aim to investigate whether missing values in the production function variables are related the firm status variables entrant, exitor and survivor. Table 19 provides information about missing values with respect to the three firm groups. The first three rows represent the shares of missing values (in %) within each firm group. That is, in the data within the group of entrants, exitors and survivors the share of missing values in the variable gross output production is given by 18.99%, 19.53% and 15.52%. The share of missing values of the variable capital is particularly high for the groups of exitors, whereas the share of missing values within the group of entrants and survivors are given by only 4.6% and 1.56%.

The last two rows show the share of entrants, exitors and survivors with respect to all observations (over firm groups), when keeping (fourth row) or dropping (fifth row) missing values. Keeping missing values, the share of entrants, exitors and survivors is given by 8.33%, 9.81% and 71.16%, respectively. Once I drop missing values, the shares of entrants and exitors, slightly drops to 7.97% and 8.12%, whereas the share of survivors increases to 73.91%. This is due to the fact that I find relatively more missing values within the groups of entrants and exitors. I generally find that missing values are related to firms' status, and, hence, productivity effects related to entry and exit might be slightly underestimated.[22]

Table 19: Missing values and firm groups^{a,b}

		Entrants	Exitors	Survivors
Share of missing values	y	18.99	19.53	15.52
	x^k	4.60	17.95	1.56
	x^l	0.00	0.00	0.00
	x^m	2.37	2.70	1.24
Share of observations	Share with NAs	8.33	9.81	71.16
	Share without NAs	7.97	8.12	73.91

[a] Figures are expressed in %
[b] Figures are calculated across all industries and years.

A4: Firms with zero or one employee

In this analysis I only consider firms with at least one employee. However, very small firms, reporting only one employee, often switch between reporting one and zero employees. Table 20

[22] Note that the shares of the three firm groups do not sum up to 100% since for the first and the last year in the sample period (1994 and 2016) I cannot identify either entrants or exitors and survivors.

shows the number of observations (across all industries) with respect to different firm size groups. As already shown in Figure 1 the evolution of the number of firms have been decreasing for all considered size groups (small, small-medium, medium-large and large). It can be seen, however that the group of small firms shows especially in 2015 a sharp decrease in the number of firms, which immediately increases in the last sample period 2016. Table 21 shows that this is due to firms shifting between zero and one employees. The very last column documents for each year (t) the number of firms that reported in $t-1$ zero employee. Here, for the year 2015, there are 4098 firms reported with zero employees that reported in 2014 one employee, inducing a sharp decrease in the number of observations (when considering firms with at least one employee). The fifth column reports analogously the number of firms that reported one employee in t and zero employee in $t-1$. It can be seen that in 2016 there are 5908 firms with one employee that reported zero employee in 2015. This leads to the sharp increase for small firms in 2016, shown in Figure 1.

Table 20: Number of observations including firms with zero employee

	Firm size groups									
# empl.	0	1	2-4	5-9	10-19	20-49	50-99	100-199	200-499	>500
1994	22526	15049	25051	13694	5483	5789	1856	1192	795	396
1995	23748	15203	25008	13664	5731	5715	1878	1201	805	393
1996	23833	15950	27213	15244	6221	5965	1906	1219	800	385
1997	24202	15707	27007	15653	6219	6040	1859	1181	809	370
1998	26350	15303	26068	14945	6573	5954	1862	1194	805	389
1999	27083	15152	25073	15204	6672	5859	1835	1197	815	404
2000	22297	14840	26376	15456	6408	5783	1872	1223	848	438
2001	20997	13972	25055	15336	6505	5805	1907	1231	896	439
2002	22051	14611	25865	16052	6903	5949	1956	1299	920	462
2003	21890	14805	25067	16285	6879	5799	1902	1273	911	455
2004	22329	14946	25296	16428	6943	5701	1858	1222	899	438
2005	20120	13726	23947	15812	6927	5542	1791	1226	865	424
2006	21626	14434	25022	16382	7417	5238	1708	1208	839	413
2007	20728	14898	25020	15995	7430	5077	1682	1150	828	406
2008	29592	13102	23683	15705	7610	5257	1717	1157	813	393
2009	31109	10334	21299	14663	7598	4875	1714	1060	756	361
2010	33011	9828	20596	14544	7556	4834	1660	1105	743	358
2011	33088	9359	19684	14310	7524	4763	1637	1048	754	334
2012	33491	8936	19197	14075	7447	4690	1628	1067	770	336
2013	36757	8126	17902	13374	7325	4627	1620	1079	735	348
2014	38298	8592	18498	13888	7584	4619	1585	1074	715	344
2015	55092	4953	10030	9386	6549	4530	1545	1047	724	350
2016	35169	13088	16992	14871	6987	4736	1545	1018	718	337

Table 21: Firms dynamics of firms reporting one and zero employee.

Year (t)	# of firms	# of firms with zero employees	# of firms with one employee	# of firms with zero employee in t-1 and one in t	# of firms with one employee in t-1 and zero in t
1994	91831	22526	15049		
1995	93346	23748	15203	1696	2111
1996	98736	23833	15950	2642	1544
1997	99047	24202	15707	1640	1729
1998	99443	26350	15303	1508	3033
1999	99294	27083	15152	2453	2959
2000	95541	22297	14840	3388	1072
2001	92143	20997	13972	1497	1521
2002	96068	22051	14611	1524	1584
2003	95266	21890	14805	1582	1555
2004	96060	22329	14946	1619	1654
2005	90380	20120	13726	1532	1541
2006	94287	21626	14434	1455	1392
2007	93214	20728	14898	1583	991
2008	101904	29592	13102	1104	3179
2009	96185	31109	10334	1218	3242
2010	98320	33011	9828	975	1976
2011	96430	33088	9359	858	1819
2012	96636	33491	8936	922	1775
2013	98149	36757	8126	887	1985
2014	101601	38298	8592	1171	1527
2015	94395	55092	4953	601	4098
2016	95565	35169	13088	5908	581

Appendix B: Production function estimation

As presented in Section 4.1 I estimate the production function by making use of the proxy variable approach, presented by Olley and Pakes (1996) and closely follow Ackerberg et al. (2015). To briefly illustrate how I proceed to estimate production I use as example a simple Cobb-Douglas gross output production function, given by

$$y_{nt} = \alpha_0 + \alpha_K x_{nt}^k + \alpha_L x_{nt}^l + \alpha_M x_{nt}^m + \omega_{nt} + \epsilon_{nt},$$

where I keep the same notation as in the main text. The first stage of the estimator consists in a non parametric estimation of the term $\tilde{\Phi}(x_{nt}^k, x_{nt}^l, x_{nt}^m)$, which was derived in equation (3), given by.

$$\begin{aligned} y_{nt} &= \alpha_0 + \alpha_K x_{nt}^k + \alpha_L x_{nt}^l + \alpha_M x_{nt}^m + \tilde{f}_t^{-1}(x_{nt}^k, x_{nt}^l, x_{nt}^m) + \epsilon_{nt} \\ &= \tilde{\Phi}(x_{nt}^k, x_{nt}^l, x_{nt}^m) + \epsilon_{nt}. \end{aligned}$$

I use the statistical software R and estimate $\tilde{\Phi}$ using the the **np** package (Hayfield and Racine, 2015). Optimal bandwidths are obtained by using the expected Kullback-Leibler cross-validation method. In the second step I regress $\widehat{\alpha_0 + \omega_{nt}}(\theta)$ on a higher order polynomial of $\widehat{\alpha_0 + \omega_{nt-1}}(\theta)$. The residuals of this regression, denoted by $\hat{\xi}_{nt}$, called the innovation to productivity, are then used to for the GMM estimation. In the case of a simple Cobb-Douglas gross output production function, the imposed moment conditions presented in equation (11) reduce to

$$E\left[\hat{\xi}_{nt}(\theta)\begin{pmatrix} x_{nt}^k \\ x_{n,t-1}^l \\ x_{n,t-1}^m \end{pmatrix}\right] = 0.$$

For the GMM regression I use the R-package **gmm** (Chaussé, 2010). The the R command **gmm()** requires first to define a function that returns a matrix where each column contains a moment conditions. I call this function "moments". In the formals of the function I define "theta", the set of parameters to be estimated, and "data", a matrix object containing all necessary variables.

```
moments <- function(theta, data){

  # Specify the production function parameters
  alphaK = theta[1]; alphaL = theta[2]; alphaM = theta[3]

  # Data
  # First step nonparametric estimate and its lagged variable
  phi_hat <- data[,"phi_hat"] ;  phi_hat_l1 <- data[,"phi_hat_l1"]

  # Explanatory variables and lagged variables
  # Capital
  k <- data[, "k"] ; k_l1 <- data[, "k_l1"]
  # Labor
  l <- data[, "l"] ; l_l1 <- data[, "l_l1"]
  # Materials
  m <- data[, "m"] ; m_l1 <- data[, "m_l1"]

  # Instruments
  z1 <- k; z2 <- l_l1; z3 <- m_l1

  # Moment matrix
  Z <- matrix(NA, nrow = nrow(data), ncol = 3)

  # Generate omega and regress on its lagged values
  omega  <- phi_hat - alphaK*k - alphaL*l- alphaM*m
  omega_lag <- phi_hat_l1 - alphaK*k_l1 - alphaL*l_l1 - alphaM*m_l1
  omega_lag_pol <- cbind(1, omega_lag, omega_lag^2, omega_lag^3)

  # Recover residuals (innovations to productivity)
  resid <- resid(lm(omega ~ omega_lag_pol))
```

```
  # Define moments
  Z[,1] = z1*resid; Z[,2] = z2*resid; Z[,3] = z3*resid

  return(Z)
}

# Specifying initial values for the numeric optimization.
res.ols = lm(y ~ k + l + m, data=data)
alphaK_0 = coefficients(res.ols)[2]
alphaL_0 = coefficients(res.ols)[3]
alphaM_0 = coefficients(res.ols)[4]

t0 = c(alphaK_0, alphaL_0, alphaM_0)

# GMM estimation
# Note: a) use "optimal" weighting matrix
#       b) use "optim" as numeric optimizer (default Nelder-Mead algo.)
res.gmm = gmm(g=moments, x=data), t0=t0, wmatrix="optimal", optfct="optim")
```

Table 22: Translog production function parameter estimates, ACF method

	Dependent variable: y_{nt}								
	2-digit Industry								
Coeff.	10 food	13 textiles	16 wood	17 pulp/paper	20 chemical products	24 metals	28 machines	29 automobiles	31 furniture
α_k	0.270***	0.196***	0.056	0.323***	0.261***	0.183**	0.306***	0.258***	0.177***
	(0.027)	(0.040)	(0.034)	(0.072)	(0.038)	(0.082)	(0.057)	(0.039)	(0.027)
α_l	0.233**	0.378	0.269**	0.353**	0.376***	0.476***	0.380***	0.462***	0.251**
	(0.102)	(0.230)	(0.128)	(0.145)	(0.145)	(0.166)	(0.126)	(0.075)	(0.121)
α_m	0.344***	0.469***	0.603***	0.403***	0.517***	0.424***	0.344***	0.327***	0.464***
	(0.044)	(0.100)	(0.068)	(0.102)	(0.095)	(0.084)	(0.069)	(0.036)	(0.075)
α_{kk}	0.068***	0.027	0.043**	−0.002	0.010	0.041	0.046	0.050**	0.021***
	(0.010)	(0.021)	(0.019)	(0.046)	(0.031)	(0.065)	(0.042)	(0.023)	(0.006)
α_{ll}	0.059	0.103	0.123	0.078	0.072**	−0.014	0.075	−0.006	0.046
	(0.062)	(0.147)	(0.106)	(0.129)	(0.031)	(0.170)	(0.106)	(0.053)	(0.130)
α_{mm}	−0.085***	0.068***	0.007	−0.061	0.030	0.049	−0.039	−0.071*	−0.104*
	(0.023)	(0.023)	(0.032)	(0.089)	(0.022)	(0.113)	(0.040)	(0.038)	(0.057)
α_{kl}	−0.130***	−0.054**	−0.057	−0.105	−0.044**	−0.055	−0.112*	−0.153***	−0.089***
	(0.026)	(0.023)	(0.043)	(0.067)	(0.019)	(0.042)	(0.060)	(0.034)	(0.027)
α_{km}	0.075***	0.003	0.006	0.070*	0.013	−0.001	0.0003	0.058***	0.038*
	(0.014)	(0.017)	(0.037)	(0.041)	(0.031)	(0.059)	(0.041)	(0.021)	(0.021)
α_{ml}	0.076***	−0.048	−0.006	0.024	−0.033	0.008	0.076	0.109***	0.095
	(0.019)	(0.063)	(0.040)	(0.118)	(0.024)	(0.095)	(0.052)	(0.040)	(0.082)
Observations	604,839	52,923	97,735	25,794	42,718	17,001	99,426	32,029	96,779

[a] *p<0.1; **p<0.05; ***p<0.01.
[b] Standard errors in parenthesis

Table 23: OLS regression of the translog production function (initial values for the GMM estimation.

	Dependent variable: y								
	10	13	16	17	20	24	28	29	31
	food	textiles	wood	pulp/paper	chemical products	metals	machines	automobiles	furniture
α_k	0.257***	0.188***	0.073***	0.317***	0.249***	0.182***	0.276***	0.296***	0.180***
	(0.004)	(0.007)	(0.006)	(0.012)	(0.010)	(0.013)	(0.007)	(0.010)	(0.006)
α_l	0.232***	0.372***	0.247***	0.349***	0.365***	0.476***	0.376***	0.466***	0.246***
	(0.006)	(0.014)	(0.011)	(0.025)	(0.017)	(0.020)	(0.012)	(0.017)	(0.010)
α_m	0.338***	0.473***	0.605***	0.399***	0.507***	0.425***	0.347***	0.307***	0.453***
	(0.005)	(0.008)	(0.008)	(0.016)	(0.012)	(0.014)	(0.008)	(0.012)	(0.007)
α_{kk}	0.059***	0.019***	0.026***	−0.004	0.002	0.041***	0.054***	0.036***	0.027***
	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.003)	(0.005)	(0.003)
α_{ll}	0.055***	0.086***	0.106***	0.078***	0.068***	−0.014	0.072***	−0.011	0.051***
	(0.005)	(0.009)	(0.010)	(0.018)	(0.011)	(0.012)	(0.008)	(0.012)	(0.009)
α_{mm}	−0.087***	0.079***	−0.060***	−0.054***	0.027***	0.049***	−0.035***	−0.085***	−0.142***
	(0.003)	(0.004)	(0.006)	(0.010)	(0.007)	(0.006)	(0.005)	(0.007)	(0.005)
α_{kl}	−0.134***	−0.053***	−0.056***	−0.115***	−0.042***	−0.055***	−0.117***	−0.138***	−0.072***
	(0.003)	(0.004)	(0.004)	(0.008)	(0.005)	(0.006)	(0.004)	(0.006)	(0.004)
α_{km}	0.080***	0.015***	0.042***	0.071***	0.021***	−0.002	0.028***	0.057***	0.038***
	(0.002)	(0.003)	(0.004)	(0.006)	(0.004)	(0.005)	(0.003)	(0.005)	(0.003)
α_{ml}	0.084***	−0.051***	0.005	0.032***	−0.032***	0.009	0.055***	0.108***	0.097***
	(0.004)	(0.005)	(0.006)	(0.010)	(0.007)	(0.008)	(0.005)	(0.008)	(0.006)
Constant	−0.019***	−0.007	0.070***	0.059***	−0.081***	0.140***	0.114***	0.060***	0.259***
	(0.005)	(0.013)	(0.008)	(0.022)	(0.018)	(0.022)	(0.011)	(0.016)	(0.007)
Observations	604,839	52,923	97,735	25,794	42,718	17,001	99,426	32,029	96,779
Adjusted R^2	0.479	0.770	0.699	0.817	0.802	0.895	0.785	0.865	0.761

Note: *p<0.1; **p<0.05; ***p<0.01

Appendix C: Productivity persistence

Estimates for the density of firms' productivity in t conditional on their productivity at $t-3$, that is $\tau = 3$. Figure 13 depicts the conditional density in a 3D plot and Figure 14 the corresponding contour plot. The results are very similar to the case presented in the main text with $\tau = 1$. Firms' show a high degree of productivity persistence.

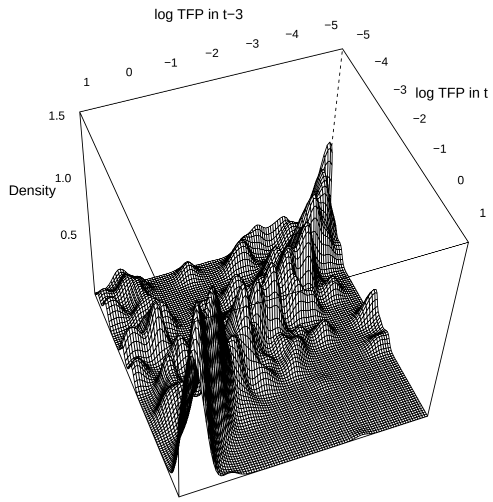

Figure 13: 3D plot of the conditional density $\hat{g}_3(\hat{\tilde{\omega}}|\hat{\omega})$

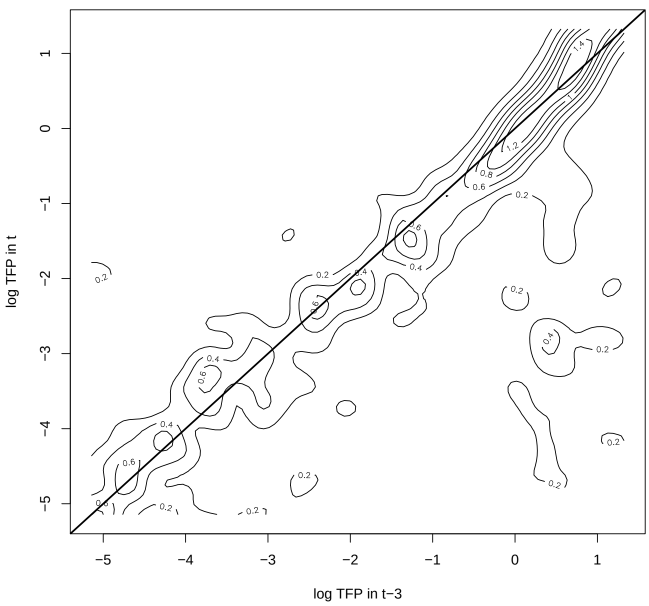

Figure 14: Contour of the conditional density $\hat{g}_3(\hat{\tilde{\omega}}|\hat{\omega})$

Appendix D: DOPD analysis 2007 - 2016

Complementary to Section 5.3, Tables 24, 25 and 26 show the industries' aggregate productivity decomposition and growth rates, however, the measures are reported for the waves 2010, 2013 and 2016, and with respect to the base year 2007.

Industries' total aggregate productivity growth

In Table 26, column(s) "total change", it can be seen that, in 2016 and relative to 2007, the industries experienced the following total change in aggregate productivity: -15.56% (food), 7.05% (textiles) 17.27% (wood), 4.46% (pulp/paper), 1.49% (chemical products), 14.13% (machines), -0.57% (metals), -35.16% (automobiles) and -7.5% (furniture). That is, four out of nine industries, namely the industries "food", "machines", "automobiles", and "furniture" show a negative aggregate productivity growth tendency.

Contribution of survivors to aggregate productivity growth

Generally, individual firms' productivity improvement (within change) of those firms that survived between 2007 and the corresponding reported wave, has positively contributed to the total aggregate productivity growth in all industries and years. With respect to 2007, the group surviving firms' increased its aggregate productivity by 8.09% (food), 12.15% (textiles), 19.95% (wood), 7.87% (pulp and paper), 9.59% (chemical products), 22.74% (metals), 4.10 % (machines), 8.21% (automobiles) and 0.77% (furniture). Instead, the contribution of the group of surviving firms to aggregate productivity growth via market share reallocation (between change) is measured to be negative for almost all years and industries, indicating that overtime market shares slowly wander from more to less productive firms. This negative contribution of the group of surviving firms by reallocation effects seems to be less important compared to individual firms productivity improvement, however, for some industries and years the amplitude is relatively important: For instance, consider the industries for "food" and "metals", here the (negative) between contribution almost completely compensates the (positive) within contribution.

Contribution of entrants to aggregate productivity growth

According to my measures and with respect to 2007, for many years and industries the group of entrants have negatively contributed to aggregate productivity growth. Here, the most significant contribution of firm entry is measured for the industries "food" and "automobiles", with a negative contribution to aggregate productivity in 2016 given by -10.76% and -88.13%, respectively. For the other industries I observe a very low contribution of entrants, which results from a very small difference in aggregate productivity between the group of entrants and survivors ($\Omega_{E2} - \Omega_{S2}$) as can be seen comparing the corresponding columns, Ω_{E2} and Ω_{S2}, in Table 26.

Contribution of exitors to aggregate productivity growth

Similar to the effect of firm entry, the contribution of firm exit is measured to be negative for nearly all industries and years, with respect to 2007. While the contribution of the group of exitors is rather small, for the industries "food" and "automobiles" the contribution of exitors to aggregate productivity growth is considerable. More precisely, with respect to 2007, in 2016 the contribution of exitors is given by -5.91% and 48.08%, respectively. Note that the contribution of exiting firms is measured by difference between survivors and exitors aggregate productivity in the base year 2007, i.e. $\Omega_{S1} - \Omega_{X1}$. A positive (negative) contribution means, hence, that the industry has lost unproductive (productive) firms (relative to survivors.

Table 24: Aggregate Productivity Decomposition (DOPD): 2007 - 2016[a]

Industry	Year	Contribution Survivors		Contribution Entrants	Contribution Exitors	Total Change[b]	Industry	Year	Contribution Survivors		Contribution Entrants	Contribution Exitors	Total Change[b]
		Within	Between						Within	Between			
10	2010	11.58	-9.08	1.85	-4.31	0.04	24	2010	6.77	-16.15	2.44	1.68	-5.26
Food	2013	8.49	-8.88	-8.38	-4.68	-13.45	Metals	2013	18.06	-19.69	0.53	7.77	6.67
	2016	8.09	-6.98	-10.76	-5.91	-15.56		2016	22.74	-16.01	-0.08	7.48	14.13
13	2010	8.09	-3.52	0.35	-3.66	1.26	28	2010	3.66	-2.29	0.16	-7.02	-5.49
Textile	2013	12.06	-4.26	-0.06	-2.00	5.74	Machines	2013	7.42	-7.58	-0.37	-0.21	-0.74
	2016	12.15	-2.94	-0.27	-1.89	7.05		2016	4.10	-2.65	-1.90	-0.12	-0.57
16	2010	16.68	-7.63	-0.05	0.53	9.53	29	2010	6.43	-40.85	1.75	-1.12	-33.79
Wood	2013	18.42	-7.84	0.85	0.31	11.74	Automobiles	2013	6.72	-3.37	-80.75	45.74	-31.66
	2016	19.95	-6.72	0.56	3.48	17.27		2016	8.21	-3.32	-88.13	48.08	-35.16
17	2010	11.58	-8.07	-0.27	-0.51	2.73	31	2010	4.24	-3.23	-0.17	-2.21	-1.37
Pulp/	2013	14.83	-7.04	-0.71	-0.68	6.40	Furniture	2013	0.20	-4.86	3.77	-2.79	-3.68
Paper	2016	7.87	-1.83	-0.41	-1.17	4.46		2016	0.77	-8.53	4.05	-3.79	-7.50
20	2010	4.20	-4.13	0.28	-1.22	-0.87	All	2010	8.14	-10.55	0.70	-1.98	-3.69
chemical	2013	10.11	-7.95	0.97	-2.36	0.77		2013	10.70	-7.94	-9.35	4.57	-2.02
Products	2016	9.59	-2.72	-1.70	-3.68	1.49		2016	10.39	-5.74	-10.96	4.72	-1.60

[a] All contributions represent growth rates in % relative to 2007.

[b] The total change in aggregate productivity is the sum of the contributions of survivors, entrants and exitors.

Table 25: Productivity Decomposition in the base year 2007 (t_1): survivors and exitors

Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_X^b	$S_X^{c,d}$	Ω_1^e	# Surv.	# Exits	Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_X^b	$S_X^{c,d}$	Ω_1^e	# Surv.	# Exits
10	2007	2010	0.05	89.67	0.47	10.33	0.10	21691	8887	24	2007	2010	-0.08	88.64	-0.23	11.36	-0.10	642	153
Food	2007	2013	0.04	81.92	0.30	18.08	0.09	16150	14168	Metals	2007	2013	-0.02	73.17	-0.31	26.83	-0.10	517	250
	2007	2016	0.03	76.16	0.28	23.84	0.09	13639	16094		2007	2016	-0.03	69.65	-0.27	30.35	-0.10	426	314
13	2007	2010	0.20	82.17	0.41	17.83	0.24	1641	622	28	2007	2010	0.02	91.15	0.81	8.85	0.09	3781	1627
Textile	2007	2013	0.22	71.89	0.29	28.11	0.24	1254	942	Machines	2007	2013	0.10	80.25	0.11	19.75	0.10	2762	2562
	2007	2016	0.21	64.32	0.27	35.68	0.23	1031	1116		2007	2016	0.10	75.13	0.10	24.87	0.10	2305	2942
16	2007	2010	-0.05	88.21	-0.10	11.79	-0.06	3528	1234	29	2007	2010	-0.66	96.61	-0.33	3.39	-0.65	1205	330
Wood	2007	2013	-0.06	78.07	-0.07	21.93	-0.06	2755	1956	Automobiles	2007	2013	-0.19	28.96	-0.84	71.04	-0.65	988	531
	2007	2016	-0.02	69.50	-0.13	30.50	-0.05	2320	2283		2007	2016	-0.17	25.98	-0.82	74.02	-0.65	820	672
17	2007	2010	-0.08	89.58	-0.03	10.42	-0.07	922	205	31	2007	2010	-0.06	84.39	0.08	15.61	-0.04	2862	2276
Pulp/	2007	2013	-0.08	80.66	-0.04	19.34	-0.07	765	339	Furniture	2007	2013	-0.07	70.06	0.03	29.94	-0.04	2016	3024
Paper	2007	2016	-0.08	74.04	-0.03	25.96	-0.07	674	411		2007	2016	-0.08	59.27	0.02	40.73	-0.04	1548	3403
20	2007	2010	0.03	87.92	0.13	12.08	0.05	1523	416	All	2007	2010	-0.07	88.70	0.13	11.30	-0.05	4199	1750
chemical	2007	2013	0.02	80.00	0.14	20.00	0.04	1268	642		2007	2013	-0.00	71.66	-0.04	28.34	-0.05	3163	2712
Products	2007	2016	-0.03	74.90	0.12	25.10	0.01	1116	764		2007	2016	-0.01	65.44	-0.05	34.56	-0.05	2653	3111

[a] According to equation (20) this table reports measures of aggregate productivity and market shares for the group of surviving and exiting firms. The measures are always (and only) for the initial year 2007.

[b] The columns Ω_S and Ω_X denote the aggregate productivity of the firm groups survivors and exitors, respectively.

[c] The columns S_S and S_X denote the aggregated market shares of the firm groups survivors and exitors, respectively.

[d] S_S and S_X are given in %.

[e] Columns Ω_1 denotes the aggregate productivity for the initial year 2007.

Table 26: Productivity decomposition in t_2: survivors and entrants[a]

Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_E^b	$S_E^{c,d}$	Ω_2^e	# Surv.	# Entr.	Industry	t_1	t_2	Ω_S^b	$S_S^{c,d}$	Ω_E^b	$S_E^{c,d}$	Ω_2^e	# Surv.	# no Entr.
10	2007	2010	0.07	89.94	0.25	10.06	0.09	21691	8117	24	2007	2010	-0.20	90.98	0.07	9.02	-0.18	642	147
Food	2007	2013	0.01	77.27	-0.36	22.73	-0.08	16150	11934	Metals	2007	2013	-0.11	78.52	-0.09	21.48	-0.11	517	209
	2007	2016	0.01	74.99	-0.42	25.01	-0.10	13639	10422		2007	2016	-0.04	70.05	-0.04	29.95	-0.04	426	164
13	2007	2010	0.23	89.88	0.26	10.12	0.23	1641	290	28	2007	2010	0.03	94.46	0.06	5.54	0.03	3781	615
Textile	2007	2013	0.25	87.87	0.24	12.13	0.25	1254	388	Machines	2007	2013	0.07	87.78	0.04	12.22	0.06	2762	838
	2007	2016	0.25	85.76	0.23	14.24	0.25	1031	341		2007	2016	0.10	78.93	0.01	21.07	0.08	2305	840
16	2007	2010	0.02	90.84	0.02	9.16	0.02	3528	816	29	2007	2010	-1.00	98.58	0.23	1.42	-0.99	1205	199
Wood	2007	2013	0.03	84.38	0.08	15.62	0.04	2755	1109	Automobiles	2007	2013	-0.17	43.97	-1.61	56.03	-0.97	988	294
	2007	2016	0.09	79.04	0.11	20.96	0.09	2320	948		2007	2016	-0.13	39.66	-1.59	60.34	-1.01	820	268
17	2007	2010	-0.05	91.56	-0.08	8.44	-0.05	922	96	31	2007	2010	-0.04	92.12	-0.07	7.88	-0.05	2862	578
Pulp/	2007	2013	-0.02	87.30	-0.07	12.70	-0.02	765	165	Furniture	2007	2013	-0.10	83.53	0.13	16.47	-0.07	2016	726
Paper	2007	2016	-0.02	78.83	-0.04	21.17	-0.03	674	178		2007	2016	-0.14	84.14	0.11	15.86	-0.10	1548	581
20	2007	2010	0.02	94.49	0.08	5.51	0.03	1523	282	All	2007	2010	-0.10	92.54	0.09	7.46	-0.10	4199	1237
chemical	2007	2013	0.02	78.23	0.06	21.77	0.03	1268	390		2007	2013	-0.00	78.76	-0.18	21.24	-0.10	3163	1783
Products	2007	2016	0.03	77.56	-0.05	22.44	0.01	1116	390		2007	2016	0.02	74.33	-0.19	25.67	-0.09	2653	1570

[a] According to equation (21) this table reports measures of aggregate productivity and market shares for the group of surviving and entering firms. The values always represent measures concerning the second time period t_2.

[b] The columns Ω_S and Ω_E denote the aggregate productivity of the firm groups survivors and entrants, respectively.

[c] The columns S_S and S_E denote the aggregated market shares of the firm groups survivors and entrants, respectively.

[d] S_S and S_E are given in %.

[e] Column Ω_2 denotes the aggregate productivity for the second time period t_2.

References

Ackerberg, D. A., Caves, K. and Frazer, G. (2015). Identification Properties of Recent Production Function Estimators, *Econometrica* **83**(6): 2411–2451.

Autor, D., David Dorn, N., Patterson, C., Booth, C. and John Van Reenen, N. (2019). The Fall of the Labor Share and the Rise of Superstar Firms, *Quarterly Journal of Economics* p. forthcoming.

Baily, M. N., Hulten, C. and Campbell, D. (1992). Productivity Dynamics in Manufacturing Plants, *Brookings Papers on Economic Activity: Microeconomics* **4**: 187–267.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013). Cross-country Differences in Productivity: The Role of Allocation and Selection, *American Economic Review* **103**(1): 305–334.

Bartelsman, E. J. and Doms, M. (2000). Understanding Productivity: Lessons from Longitudinal Microdata, *Journal of Economic Literature* **38**(3): 569–594.

Bernard, A. B. and Jensen, J. B. (1999). Exporting and productivity, *National bureau of economic research* (w7135).

Blanchard, P., Huiban, J. P. and Mathieu, C. (2014). The Shadow of Death Model Revisited with an Application to French Firms, *Applied Economics* **46**(16): 1883–1893.

Chaussé, P. (2010). Computing Generalized Method of Moments and Generalized Empirical Likelihood with R, *Journal of Statistical Software* **34**(11): 1–35.

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-level Export Status, *American Economic Review* **102**(6): 2437–2471.

De Monte, E. and Koebel, B. (2020). Entry and Exit in Cournot Equilibrium with Heterogeneous Firms, *mimeo, Beta, Université de Strasbourg* .

Ericson, R. and Pakes, A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work, *The Review of Economic Studies* **62**(1): 53–82.

Esponda, I. and Pouzo, D. (2019). The Industry Supply Function and the Long-run Competitive Equilibrium with Heterogeneous Firms, *Journal of Economic Theory* **184**: 104946.

Fariñas, J. C. and Ruano, S. (2005). Firm Productivity, Heterogeneity, Sunk Costs and Market Selection, *International Journal of Industrial Organization* **23**(7-8): 505–534.

Foster, L., Haltiwanger, J. C. and Krizan, C. J. (2001). Aggregate Productivity Growth. Lessons from Microeconomic Evidence, *New Developments in Productivity Analysis*, University of Chicago Press, pp. 303 – 372.

Foster, L., Haltiwanger, J. and Krizan, C. J. (2006). Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s, *Review of Economics and Statistics* **88**(4): 748–758.

Foster, L., Haltiwanger, J. and Syverson, C. (2008). Reallocation, Firm turnover, and Efficiency: Selection on Productivity or Profitability?, *American Economic Review* **98**(1): 394–425.

Geroski, P. A. (1995). What Do We Know About Entry?, *International Journal of Industrial Organization* **13**(4): 421–440.

Griliches, Z. and Regev, H. (1995). Productivity and Firm turnover in Israeli Industry: 1979-1988, *Journal of Econometrics* **65**: 175–203.

Haltiwanger, J. (2011). Firm Dynamics and Productivity Growth, *European Investment Bank Papers* **16**(1): 116–136.

Hayfield, T. and Racine, J. S. (2015). Nonparametric Econometrics: The np Package, *Journal of Statistical Software* **27**(5).

Henderson, D. J. and Parmeter, C. F. (2015). *Applied Nonparametric Econometrics*, Cambridge University Press.

Hopenhayn, H. A. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium, *Econometrica* **60**(5): 1127–1150.

Johnson, P. A. (2000). A Nonparametric Analysis of Income Convergence across the US States, *Economics Letters* **69**: 219–223.

Johnson, P. A. (2005). A Continuous State Space Approach to "Convergence by Parts", *Economics Letters* **86**: 317–321.

Jovanovic, B. (1982). Selection and the Evolution of Industry, *Econometrica* pp. 649–670.

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione, *Giorn. Inst. Ital. Attuari* **4**: 83–91.

Leung, D., Meh, C. and Terajima, Y. (2008). Firm size and productivity, *Bank of Canada Working Paper* (45).

Levinsohn, J. and Petrin, A. (1999). When industries become more productive, do firms?, *National Bureau of Economic Research* (w6893).

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables, *The Review of Economic Studies Review of Economic Studies* **70**(2): 317–341.

Li, Q. and Racine, J. S. (2011). *Nonparametric econometrics: Theory and practice*, Princeton University Press.

Melitz, M. J. and Polanec, S. (2015). Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit, *RAND Journal of Economics* **46**(2): 362–375.

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry, *Econometrica* **64**(6): 1263–1297.

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants, *Review of Economic Studies* **69**: 245–276.

Polanec, S. (2004). On the evolution of size and productivity in transition: Evidence from slovenian manufacturing firms, *LICOS Discussion Paper* (154).

Smirnov, N. V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples, *Bull. Math. Univ. Moscou* **2**(2): 3–14.

Syverson, C. (2004). Product Subsitutability and Productivity Dispersion, *Review of Economics and Statistics* **86**(2): 534–550.

Syverson, C. (2011). What Determines Productivity?, *Journal of Economic Literature* **49**(2): 326–65.

Van Biesebroeck, J. (2005). Firm Size Matters: Growth and Productivity Growth in African Manufacturing, *Economic Development and Cultural Change* **53**(3): 545–583.

Van Biesebroeck, J. (2008). The Sensitivity of Productivity Estimates, *Journal of Business & Economic Statistics* **26**(3): 311–328.

Wagner, J. (2010). Entry, Exit and Productivity: Empirical Results for German Manufacturing Industries, *German Economic Review* **11**(1): 78–85.

Wooldridge, J. M. (2009). On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables, *Economics Letters* **104**(3): 112–114.