KU LEUVEN

CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS14.25

AUGUST 2014

Job creation, firm creation, and *de novo* entry

Karen GEURTS and
Johannes VAN BIESEBROECK

Econometrics

Faculty of Economics
And Business

Job creation, firm creation, and *de novo* entry*

Karen Geurts   Johannes Van Biesebroeck§

August 2014

Abstract

Firm turnover and growth recorded in administrative data sets differ from underlying firm dynamics. By tracing the employment history of the workforce of new and disappearing administrative firm identifiers, we can accurately identify *de novo* entrants and true economic *exits*, even when firms change identifier, merge, or split-up. For a well-defined group of new firms entering the Belgian economy between 2004 and 2011, we find highly regular post-entry employment dynamics in spite of the volatile macroeconomic environment. Exit rates decrease with age and size. Surviving entrants record high employment growth that is monotonically decreasing with age in every size class. Most remarkably, we find that Gibrat's law is violated for very young firms. Conditional on age, the relationship between employment growth and current size is strongly and robustly positive. This pattern is obscured, or even reversed, when administrative entrants and exits are taken at face value. De novo entrants' contribution to job creation is relatively small and not very persistent, in particular for (the large majority of) new firms that enter with fewer than five employees.

* The authors are grateful to the Belgian National Social Security Office (NSSO), in particular to Peter Vets, for providing the data and for helping to develop the employee flow record linking method; and to Statistics Belgium, in particular to Youri Baeyens and Antonio Fiordaliso, for linking the NSSO data to the Belgian Business Register and for providing firm record linkages created for the Eurostat Structural Business Statistics. Financial support from ERC grant No. 241127 and KU Leuven project financing is gratefully acknowledged.

§ Geurts: University of Leuven (KU Leuven), Karen.Geurts@kuleuven.be; Van Biesebroeck : University of Leuven (KU Leuven) and CEPR, Jo.VanBiesebroeck@kuleuven.be.

1 Introduction

Firm-level employment dynamics cannot be viewed separately from firm dynamics. Firms enter and eventually exit from industry, but in between they can be restructured or change legal identifier for a variety of reasons. Simply looking at the group of entrants and exits in an administrative sense, tracing appearances and disappearances of firm identifiers in the national firm registry, gives a highly misleading impression of the growth patterns of young firms.

Focusing on a well-defined group of *de novo* entrants and controlling for various changes in identifier, mergers or split-ups, we find highly regular patterns of post-entry dynamics. Larger young firms have lower exit rates than smaller ones and, remarkably, also exhibit higher growth rates. This leads to a positive growth-by-size relation in the first years after entry, a feature that contrasts with most earlier literature. The passive learning model of Jovanovic (1982), where firms expand or exit the market as they receive imperfect information about their own quality from their market operations, can rationalize these early growth patterns.

Since large-scale business registers have come into use, researchers have recognized that many entries and exits are the result of changes in identifier or firm structure of continuing firms. Such breaks in the longitudinal linking of individual firm records naturally introduce an upward bias in measures of firm and employment dynamics (Spletzer, 2000). We show that it also affects the growth-size relationship of young firms and their contribution to job creation, both at the time of entry and in subsequent years.

We take great care to as much as possible only identify firms as de novo entrants when they start new operations, corresponding to firm creation in theoretical models of firm dynamics (Jovanovic, 1982; Hopenhayn, 1992). Working with the universe of firm-year observations for Belgian employers from 2003 to 2012, we use two sets of longitudinal firm linkages to distinguish de novo entrants and true exits from firms that operate continuously but change identifier. The first is developed by Statistics Belgium in line with the OECD-Eurostat recommendations on business demography statistics (Eurostat-OECD, 2007). It relies on additional administrative data sources and probabilistic matching. We supplement these official links with an alternative method based on employee-flow information. Tracing the employment history of clusters of workers, one key input that can be followed unambiguously over time, we can link additional firm records and bridge changes in identifier and firm structure. This approach refines the method pioneered for Canada by Baldwin, Dupuy and Penner (1992) and used for the United States by Benedetto et al. (2008).

In the sample of de novo entrants, several patterns of post-entry dynamics are consistent with the consensus in the literature. First, conditional on age, the survival rate increases with firm size and, conditional on size, it increases with firm age—as in Dunne, Roberts and Samuelson (1989) or Mata, Portugal and Guimaraes (1995). After carefully distinguishing true economic exit from changes in firm identifier, both patterns are remarkably regular, even over the volatile sample period that includes the 2008-2009 recession. Second, growth rates of surviving young firms monotonically decline with firm age within every size class—as in Evans (1987b) and Decker et al. (2014).

One pattern, however, is in sharp contrast with most earlier literature which has found a negative relationship between growth and size conditional on age for surviving young firms (Evans, 1987a, 1987b; Mata, 1994; Lotti, Santarelli and Vivarelli, 2003).[1] We instead find that in each entry cohort larger firms tend to increase employment more quickly than smaller firms. As a result, growth rates of de novo entrants are increasing in firm size conditional on age, confirming recent findings by Haltiwanger, Jarmin and Miranda (2013) and Decker et al. (2014). We find that the positive growth-size relationship is particularly pronounced in the first years after entry, but converges to more proportional growth rates by age five. Gibrat's law, which posits that firm growth is independent of size (Sutton, 1997), is strongly violated in the first years after entry and this remains true in several robustness checks.[2]

To some extent, our finding of a reversed pattern reflects a difference in focus. First, we only look at the initial years after entry and illustrate that the increasing pattern disappears as firms mature. Most previous studies study the entire population of firms or use a very coarse age classification. In some cases, growth rates are estimated conditional on surviving for several years which excludes a large share of small firms that exit early and introduces an upward bias in the growth rate of small entrants. Second, our sample includes the smallest entrants, even single-employee firms, while most studies have a minimum size threshold of five or even more employees. This matters greatly as the increasing pattern is concentrated in firms ranging from one to twenty employees. Hsieh and Klenow (2014) are careful to identify the universe of firms and they also find remarkably little dynamism for small firms in India. Third, our sample contains all private sector firms, while most earlier studies focus on manufacturing where the increasing pattern is also less pronounced in our data.[3]

Finally, and most importantly, data sets used in previous studies do not, or only partially correct for missing links in firm histories due to ownership changes, mergers, splits or other forms of identifier changes. We exclude so-called *de alio* entrants which continue the activities of a pre-existing firm under a new identifier. On average these are much larger than de novo entrants and have much lower exit and growth rates, making them almost indistinguishable from incumbents.[4] As the official link method fails to detect many of the larger de alio entrants, only 1 percent of all administrative entrants but representing 26

[1] The negative relationship also holds in the pooled age class of firms younger than five years in Dunne, Roberts and Samuelson (1989).

[2] It also contrasts with the received wisdom among policy makers that "small firms are the engine of job creation." Previous work already qualified this attribution by highlighting the correlation between age and size and crediting young firms especially with most job creation (e.g. see Haltiwanger, Jarmin and Miranda, 2013).

[3] We do exclude personal and social services, a sector that in Belgium is dominated by quasi-government entities, and a special sector dominated by a highly subsidized voucher system.

[4] Most administrative entrants in larger size classes are de alio, e.g. 60% of firms in the 10-19 employees size class, even 90% in the 50+ class. Including them lowers the average growth rate of larger entrants substantially.

percent of their employment, the worker flow method is crucial to correctly identify all firm histories.

When a firm identifier disappears, we can make a similar distinction between true economic *exit* and firms merely *transferring* their activities to existing or new legal entities. Such transfers do not seem to be random events as they are more likely for larger and more successful entrants. Taking care not to classify them as firm exits, we obtain more reliable estimates of exit rates, but also of organic growth rates.

Both de alio entry and firm transfers are important economic phenomena that are likely to contribute to economic dynamism.[5] Most ID changes, however, are transformations of established firms for legal, liability, tax or administrative reasons. They have little to do with economic entry and exit as the firm continues the same activities with the same work force. As they disproportionately involve larger firms, lumping low-growth de alio entrants with de novo entrants and misclassifying transfers as exit induces a strong downward bias in the growth-size relationship.

The positive growth-size relationship we document is consistent with recent findings for the United States.[6] The primary contribution of Haltiwanger, Jarmin and Miranda (2013) is to show that young, rather than small firms create most jobs. Young firms just tend to be small on average. Conditional on firm age, they also find a (weak) positive relationship between employment growth and size. More recently, Decker et al. (2014) show for the same dataset that the positive relationship holds in every age category. Both studies work with the universe of U.S. firms, including firms of all sizes and all sectors, as we do, and they pay specific attention to accurately determine firm age.[7] They interpret the patterns as evidence of an “up or out” dynamic. A subset of high-growth firms has higher survival rates and persistently higher growth rates. The presence of these firms gives each entry cohort a lasting positive impact on job creation.

Our findings differs in two ways. First, in our sample the positive growth-size relationship is stronger than theirs in the first years following entry, but becomes weaker as entry cohorts mature. It disappears almost entirely by the time firms have survived approximately five years.[8] Second, de novo entrants in the Belgian sample span only a very limited employment range at start-up. Firms that employ fewer than five workers account on average for 70

---

[5] Some studies, e.g. Eriksson and Kuhn (2006) and Muendler, Rauch and Tocoian (2012), show specific post-entry dynamics for spin-offs, which are included in the set of de alio entrants we do not focus on.

[6] Mata and Portugal (2004) also show evidence consistent with our finding. They find that, conditional on age, new plants started by foreign firms grow more rapidly than plants of domestic firms, even though the former are on average six to seven times larger.

[7] To identify firms that operate continuously, and by extension entry and exit, the data collection for the U.S. Census and Annual Survey of Manufactures follows economic activity by establishment address rather than by legal entity.

[8] Table 1 in Decker et al. (2014) shows growth rates to be increasing by size-class for every single age group, ranging from 1-2 years to 16+ in two-year increments. The positive relationship even holds for (survivors of) very mature cohorts.

percent of an entry cohort's workforce, while only 2 percent of employment is in firms that start with more than fifty workers.[9]

With these differences, the patterns we find are very much in line with the passive learning model of Jovanovic (1982). Firms are assumed to enter without knowing their innate productivity. Young firms are still highly uncertain about their own quality and respond to market success by expanding. As firms mature and learn their productivity, they converge to their steady state size.[10] The inverse relationship between exit and age, as well as between exit and size within age cohorts are consistent with a noisy selection process. Unsuccessful firms stay small and eventually choose to exit. Larger firms are the ones that have received favorable cost information in previous periods and have expanded. Subsequent information becomes gradually less informative and is less likely to induce exit.

The positive relationship between growth and size in early years is consistent with the same mechanism if young firms cannot immediately operate at their optimal size. While a firm's estimate of its own productivity might be very high, credit and regulatory constraints or adjustment costs can limit growth in the first years. For some expanding firms, current size will be below desired size. If expansion is only gradual, larger firms will continue to grow relatively fast for some years. Their size and growth rate both reflect underlying firm quality. While this process unfolds in continuous time, we always observe new firms entering our data set on June 30. Larger entrants are firms for which this process has already been unfolding over the past year and they have already reached a certain size upon entering the data set. In the years immediately following entry, larger firms keep growing more rapid than average, until current size catches up with desired size and the positive relationship between growth and size disappears.

The contribution of entrants to job creation depends on their absolute importance, reflected in the initial firm size distribution, and on changes in the distribution over time, captured by exit and growth patterns. Each of these dimensions is also influenced by the *de novo* versus *de alio* and the *exit* versus *transfer* distinction. After the necessary corrections, we find that the initial contribution of de novo entrants to job creation in the Belgian economy is very small.[11] Averaging over the nine year sample period, more than two thirds of all jobs created by new entrants are in firms with at most four employees and only one job in seven is in firms employing at least ten workers at startup. This pattern might be intuitive, but contrast with much of the literature that is based on datasets with a minimum employment threshold. They also contrast with OECD statistics suggesting that in half of the member countries one in twenty five active firms with at least 10 employees is a new entrant (OECD, 2013).

---

[9] The corresponding statistics from Table 1 in Haltiwanger et al. (2013) are 21 and 35 percent. In the United States, 13 percent of an entry cohort's employment is in firms that start out with at least 500 workers in their very first year. It seems highly likely that some of this entry represents a continuation of existing activities.

[10] In the related model of Hopenhayn (1992), even mature firms experience random productivity shocks that induce random growth rates in steady state, but these are unrelated to firm size.

[11] Only 9% of new firm identifiers are identified as de alio entrants, but they account for almost half of the employment in new identifiers and many of them are not detected by the official data linking methods.

Finally, we show that firm size at startup has some predictive power for subsequent job creation. For the group of small entrants, job destruction associated with firm exit outweighs job creation through internal growth from the start. As they dominate the entry cohort, it limits the contribution of new entrants to job creation. Employers that enter with at least five employees are better able to maintain or even increase the initial employment level. Only after five years does the total employment of the group fall below its entry level. It suggests that the lack of dynamism that Hsieh and Klenow (2014) document in small firms in India and Mexico, is not limited to developing countries.

The remainder of the paper is organized as follows. We first describe the data in Section 2. This is followed in Section 3 by a discussion of record linking using the official and the employee-flow methods and how it affects the identification of entry and exit. In Section 4 we describe the patterns of survival and growth rates by age and size and in Section 5 we summarize the implications for overall job creation by entrants. Section 6 concludes.

2 Data

The analysis is based on a firm-level data set maintained by the Belgian National Social Security Office (NSSO). It covers the universe of firms with at least one employee in Belgium. Firms are identified by their official Belgian enterprise number, the CBE number (Crossroad Bank for Enterprises), which is a unique firm identifier used by all government administrations. New enterprises receive this identifier upon registration and keep it for their entire lifetime, also when their legal status or ownership changes. In contrast to some other countries, administratively induced changes to an enterprise number are very rare. As a result, administrative entrants are readily identified as new CBE numbers entering the NSSO database, which happens the first year they record positive employment.

The NSSO data set is based on quarterly declarations of employers to the Social Security Office about their employees. Employees themselves are also identified by a unique personal identification number.[12] We will exploit the linked employer-employee information in the NSSO data to trace the continuity of firms over various changes in identifier, mergers or split-ups.

We restrict our analysis to a subset of the NSSO database. For comparability with other research, only firms in the private, non-farm sector are included. We also exclude highly subsidized sectors where firm and job creation receive strong support from government programs and which grew strongly over the sample period. Excluded are “human health and social work activities”, a sector in which most expenditures are publicly financed. We also exclude sectors that make use of service vouchers, a scheme for household help established in 2004, which subsidizes 70% of the wage cost.[13] Since price competition in these activities is

[12] The personal identification numbers are from the National Register and are assigned at birth.

[13] Employment in Health and social work increased by 39% between 2003 and 2012 (from 9% to 11% of total salaried employment); in the same period, the number of employees in the service voucher system increased from 0% to 4% of total salaried employment.

weak or absent, firm performance and dynamics are likely to be different from the competitive markets we want to study.[14]

The analysis covers the period 2003-2012 which allows identification of entrants from 2004 onwards. The data includes an average of 178,000 active firms and 2,070,000 employees per year. Aggregate employment in this set of firms increased by 0.9 percent per year till 2008, dropped by 2.5% between 2008 and 2010 and has been more or less stable since then.

3 Linking firms

Between the moment a firm starts up operations and exits the market, its administrative identifier may change for three reasons. In some countries, a firm is administratively assigned a new identification number when its ownership or legal status changes. Such changes are exceptional in Belgium, which reduces the incidence of spurious entry and exit.[15] Firms may induce an ID change themselves for tax or liability reasons, closing down the existing business number and continuing the same activities in a newly registered company. Continuing firms may also receive a new ID or newly formed firms may use an existing ID following a merger, acquisition, or spin-off.

Such events hamper the analysis of firm and employment dynamics in several ways. Misinterpreting changes in administrative identifiers as the closing and start-up of firms yields an overestimation of entry and exit, and the associated job creation and destruction. It introduces a bias in the growth-by-size or exit-by-size relationships, as the probability of being involved in an event often depends on size. It also hampers comparative analysis, because the incidence of changes in firm identifiers depends on country-specific administrative, legal and tax regulations.[16]

Our objective is to identify firms as de novo entrants in a way that corresponds to firm creation in theoretical models of firm dynamics (Jovanovic, 1982; Hopenhayn, 1992). As much as possible, we only identify firms as new entrants when they satisfy the following definition from the Eurostat-OECD (2007) *Manual on Business Demography Statistics*:

> *"The birth of an enterprise amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event. To be excluded are (i) enterprises created as the result of a break-up, merger, split-off,…, (b) new enterprises*

[14] Table A.1 in the Appendix lists the sectors included in the analysis, as well as their classification into six main industries that we use throughout the paper.

[15] In some countries, the official ID even changes when a firm switches accountants. Vilhuber (2009) provides an overview and Baldwin et al. (1992), Jarmin and Miranda (2002), and Hethey and Schmieder (2013) provide details, respectively, for Canada, the United States, and Germany. In Belgium, the CBE number is not affected by changes in ownership or legal status, except for the transition of self-employed activities into a legal company.

[16] These issues are further discussed in Spletzer (2000), Baldwin, Beckstead and Girard (2002), Benedetto et al. (2007), and Bartelsman, Haltiwanger and Scarpetta (2009).

that simply take over the activity of an existing firm, (c) additional legal units solely created to provide a single business function of the company."

One way to identify firms that operate continuously, and by extension entry and exit, is to follow economic activity by address rather than by legal entity. The data collection for the U.S. Census and Annual Survey of Manufactures follows this approach. Establishment-year observations are linked over time if an economic activity continues at the same address, irrespective of which firm owns the establishment. Haltiwanger, Jarmin and Miranda (2013) use this information to highlight the importance of young firms for job creation. Researchers in developing countries have taken this approach one step further by conducting exhaustive censuses of firm activities in well-defined geographic areas and asking recall questions to construct past growth rates. Mead (1994) and Hsieh and Klenow (2014) followed this approach in sub-Saharan African countries and India.

Working with the universe of administrative firm-year observations for Belgium, we use an alternative, the employee flow method to identify firms that operate continuously but change identifier. Tracing the employment history of individual workers, one key input that can be followed unambiguously over time, we can verify whether new administrative entrants are truly new firms. We refine the method pioneered for Canada by Baldwin, Dupuy and Penner (1992)[17] and used for the United States by Benedetto et al. (2008). The method used in this paper has been developed in collaboration with NSSO (Geurts and Vets, 2013).

In addition, our data set incorporates an extensive set of firm-level linkages developed by Statistics Belgium in accordance with the Eurostat (2003) and Eurostat/OECD (2007) guidelines for constructing indicators on firm dynamics. They exploit information from a comprehensive database that combines firm-level information from different administrations such as the national register of legal entities, the trade register, VAT declarations, and Social Security reports.[18] They also use probabilistic matching, based on firm name, address, and industry code, to reclassify some administrative entrants as incumbents and some exits as continuing firm.[19] We denote these links as 'identified by the official method.' We can validate them and identify additional links by the employee flow method.

3.1 The employee flow method

Employee flow methods have been developed to produce more reliable statistics on firm and employment dynamics. By removing spurious firm entries and exits, they eliminate spurious job creation and destruction resulting from administrative reshuffling of employment between business numbers (Albaek and Sorensen, 1998; Baldwin, Beckstead and Girard, 2002). They are also used to identify changes in firm structure such as mergers and acquisitions

[17] Since 1992, the method is used to construct firm linkages in the Canadian National Accounts Longitudinal Microdata File (NALMF).

[18] DBRIS 2 is the enterprise data warehouse developed by Statistics Belgium which incorporates information from administrative registers and business surveys.

[19] The probabilistic matching procedure is based on the Term Frequency - Inverse Document Frequency method. Results are extensively checked by automatic rules and manual controls.

(Mikkelson, Unger and LeBel, 2006), or to examine a specific group of entrants such as spin-offs (Eriksson and Kuhn, 2006; Muendler, Rauch and Tocoian 2012). To our knowledge, they have not been used to trace firm histories from the date of entry over potentially multiple changes in firm identifiers. This exercise cannot be extended too far as some entrants are involved in successive firm ID changes, mergers or spin-offs. Following the employment growth of entrants over multiple events rapidly becomes too complex and dependent on assumptions. We therefore restrict our analysis to the first five years after start-up.

To establish the longitudinal history of a firm, we follow one of its main production factors, the work force. If a firm identifier changes but the firm continues its activities, the workforce of the old and new firm ID will largely be the same. Continuity of the workforce is thus used to identify continuity of the firm. If a sufficiently large cluster of employees 'moves' between two identifiers in a short period of time, it is unlikely to be the result of inter-firm worker mobility but rather an indication of a change, merger or split-up of firm identifiers. We identify such events and do not classify firms that start this way as de novo, but rather as de alio entrants and firms that leave this way as transfers, rather than exits.

Employee flow methods have been implemented using a minimum cluster of three to five employees between firm IDs.[20] For smaller clusters, there is a high probability that employee flows merely represent individual worker mobility. As a result, the method cannot identify restructuring of small firms and we apply it in addition to the official record linking method which does capture links between small firms.

The absolute cluster size is supplemented with a set of relative thresholds. Five employees leaving one firm ID for another will mean something different if the old employer had a workforce of 7 or 27. The exact levels are not critical since the distribution of the relative cluster sizes is strongly U-shaped. Most clusters represent either a very small share of the firm's workforce, indicating individual worker mobility, or a very large share, indicating a change in firm identifier.

The key decision rules to establish a link between different identifiers of the same firm are as follows. First, the clustered employee flow has to be observed between two successive quarterly observations $q-1$ and q. This short time span makes the movement of a sizable employee cluster unlikely to be the result of individual worker mobility. Second, only clusters of at least five employees are taken into account. The third rule sets the relative threshold of the cluster: a link is established between two firm IDs if the employee cluster represents at least 50 per cent of the work force of both the origin firm in $q-1$ (the 'predecessor'), and the destination firm in q (the 'successor'). The objective of this threshold is to only connect firm IDs with mostly the same workforce. 78 percent of the employee flow links we identify at entry meet this key rule.

A set of additional rules are added to also identify links between firms following mergers or spin-offs. A common example is a predecessor that transfers only part of its activities to a newly created business number. If this part is smaller than half of the predecessor's work

[20] See Benedetto et al. (2007), Hethey and Schmieder (2013) and Rollin (2013) for examples.

force, the earlier 50 percent rule will not capture the event. The workforce of the new business number, however, will predominantly consist of employees that are transferred from one predecessor. To identify such spin-offs, we impose that the clustered employee flow represents at least 75 per cent of the workforce of the new identifier. Another 18 per cent of the employee flow links that we identify at entry meet this rule. Similar decision rules capture other forms of organizational restructurings, but they represent many fewer links. The full set of rules is presented in Table A.2 in the Appendix.

3.2 Identifying de novo entry

The links established by both the official and the employee flow method are first used to identify administrative entrants that do not start *de novo*. These firms are the result of a relocation of existing production factors from an incumbent, either in total or partially, to a new business number. We denote them as *de alio* entrants, i.e. emerging from an established firm. The distinction between de novo and de alio entrants exploits quarterly information on business numbers entering and leaving the NSSO data set. Yet all summary statistics and post-entry firm dynamics will be based on employment measured at a fixed annual point of observation, namely the end of the second quarter of each year.

Results in Table 1 show that, in the year of entry, de alio entrants account for 9 percent of all administrative entrants. The probability that a new business number corresponds to de novo entry of a new firm decreases dramatically with size. More than one third of administrative entrants with 5 to 9 employees and two thirds of those with 10 or more employees are identified as de alio. De novo entrants with 50 or more employees are extremely rare.

Table 1 further shows that the two methods are complementary.[21] The official method is especially needed in the size class below five employees, where employee flow links are absent by construction. Yet the employee flow method identifies two to three times more de alio entrants in larger size classes—the vast majority result from simple ID changes (78%, see Table A.2 in the Appendix). Another sizeable fraction arises from spin-offs of an incumbent (18%). By definition, de alio entrants identified by an employee flow link have at least 50 percent of their workforce in common with a predecessor. In practice, this share is much higher for most de alio entrants, with a median of 94 percent and a mean of 90 percent.

— Table 1 approximately here —

Although a small group, de alio entrants introduce a strong bias in entry statistics based on administrative data due to their size. They exhibit incumbent-like features which obscure the distinct characteristics of de novo entrants. Figure 1 illustrates the pronounced difference between the employment distribution of all administrative entrants versus that of de novo entrants. Comparing the black and grey bars highlights that employment of de novo entrants

[21] The fractions identified by the two methods sum to more than the total because some links are identified by both methods.

is highly concentrated in the first two size categories (70%), while half of the employment of de alio entrants is in the top two size categories.

It especially introduces a bias in measures of job creation by new entrants. While de alio entrants represent only 9 percent of administrative entrants, they account for almost half of their employment (44%).[22] As they enter through administrative transfers of existing jobs to a new firm identifier, their employment should not be considered new job creation. Our narrow identification of de novo entrants suggests that the magnitude of job creation through entry is very modest. A large number of new firms enter the market de novo every year (8.4% of all active employers), but the number of jobs they create at start-up represents only 1.5% of total private sector employment.

Another indication that de alio entry differs from our usual understanding of entry is in their relation to the business cycle. Business formation is considered to be procyclical, and employment-weighted entry is found to covary positively with output growth (Campbell, 1998). Our de novo entrants strongly reflect this feature as annual changes in their job creation show a strong positive correlation of 0.81 with GDP growth. This relationship is almost entirely absent for job creation by de alio entrants which shows hardly any correlation (0.13) with GDP growth. It suggests their entry is largely motivated by legal, tax or administrative reasons and to a lesser extent by economic events.

— Figure 1 approximately here —

The official Eurostat/OECD method helps to avoid these biases somewhat, but it fails to identify many de alio entrants. The right panel in Figure 1 illustrates that even though the official method identifies the largest number of de alio entrants, they are predominantly small and do not account for a significant fraction of employment. Moreover, virtually all de alio entrants it identifies with a workforce of at least 10 employees are also picked up by the worker flow method. In contrast, more than half of all employment by de alio entrants with at least 5 employees is only identified using worker flows and not by the official method. If only one method is used, the employee flow method is clearly preferable.

Summary statistics in Table 2 underscore the importance of distinguishing between the two types of entrants and their difference. The administrative and official method locate a significant share of entrants' employment in larger firms, and assign an average entry size of 3.1 and 2.7 employees respectively. De novo entrants, by contrast, are much smaller (1.9 employees) and exhibit a strongly right-skewed employment distribution. This is already the case if only the employee flow method were used. De alio entrants are much larger on average (14.6 employees) and the mass of employment is located in larger firms. Both features resemble those of incumbents. These patterns appear in all sectors, but are especially pronounced in manufacturing: only 27 percent of administrative job creation is with de novo entrants and 66 percent of de alio entrants start with at least 50 employees.

---

[22] In some countries, but not in Belgium, changes in ownership or legal structure often lead to new business numbers as well. The bias in job creation statistics is likely to be even higher in such countries.

— Table 2 approximately here —

Not distinguishing between both groups of entrants mixes up two distinctly different populations. On the one hand, de novo entrants are newly created firms and typically very small. On the other hand, de alio entrants, like incumbents, have been created (much) earlier, have already survived for some years and are typically much larger. The overall employment distribution of de alio entrants is almost as skewed towards larger size categories as the employment distribution of incumbents—shown with a dashed line in Figure 1.[23] Properly distinguishing between the two groups matters greatly for growth and exit patterns, as we illustrate below.

3.3 Identifying economic exit

We already described how to identify de novo entrants. Subsequently, these firms may disappear from the sample because they change identifier or merge with another firm. Miscoding such events as exits leads to an overestimation of failure rates. As it particularly involves larger firms, it also biases the exit-size relationship. Events that relocate activities between business numbers will be denoted as *transfers* and distinguished from true economic *exit*. We again use both the official and the employee flow record links to identify firms involved in a transfer.

Such transfers may also affect surviving firms, for example when they shift only part of their activities and associated workforce to a newly created business number. Misinterpreting the accompanying employment drop as job destruction induces a negative bias in the firm-level growth rate. Employment of young firms involved in a transfer, both if the firm identifier disappears or continues, is imputed beyond the event by making use of employment of the linked successors. This is discussed further in the next section.

In the first five years after start-up, the business number of half the de novo entrants is discontinued. Statistics in Table 3 illustrate that only 4 percent of these are identified as transfers, but the fraction is much higher for larger firms. Approximately 30 percent of young firms with ten or more employees in the year before their ID is discontinued, do not exit the market. They continue with a different ID and employ mostly the same workforce. The employee flow links show that again most transfers (77%) correspond to straightforward ID changes; 19 percent are absorbed by another business number; and 4 percent, especially in larger size classes, are split into several business numbers.[24]

— Table 3 approximately here —

Figure 2 illustrates the importance of distinguishing transfers from exits to estimate the relationship between the exit probability and firm size of de novo entrants. The black areas

[23] Including employment in the subsidized sectors, see the discussion in the data section, leads to an employment distribution for de alio entrants that is almost indistinguishable from that of incumbents.

[24] Table A.2 in the Appendix lists the different rules with relative thresholds to identify mergers and spin-offs involving multiple business numbers. The average transfer involves 86% of a firm's workforce.

represent employment of firms that exit; the grey areas employment of firms that continue with a different ID. Almost all de novo entrants with fewer than 5 employees that disappear from the sample are economic exits. However, for the other size classes, a substantial share of employment of discontinued firm IDs is transferred to a new business number. As a result, the overall employment distribution by firm size of exiting firms is even more right-skewed if transfers are classified correctly as continuing firms. Many small firms exit and this mostly reflects job destruction. Many larger firms also disappear, but a significant share of these are mere transfers of activities to other entities and involve no, or limited job destruction.[25]

— Figure 2 approximately here —

Figure A.1 in the Appendix highlights that firms leaving the sample through exit or through transfer tend to be very different. It shows the average size of de novo entrants conditional on age broken into three groups. Firms that exit are always the smallest, ranging from an average of 1.5 to 2.5 employees depending on age. Firms that continue unchanged tend to be larger, averaging 2 to 3 employees. Firms involved in a transfer are up to four times larger than exiting firms and 2 to 3 times larger than continuing firms.

3.4 Identifying organic growth

To impute the number of jobs in the years after a transfer, we link disappearing identifiers to their successors. In case of one-to-one ID changes, the vast majority of events, imputed employment is simply total employment of the successor. The same holds for spin-offs or break-ups. The imputed number of jobs in the subsequent periods equals the sum of jobs of the different successors. Only when a firm merges with another firm can its future employment level not be observed directly. In that case, we make the assumption that the average contribution of the entrant to the employment growth of the merged entity is proportional to its employment share at the time of the merger. Therefore, the imputed employment of the original firm following a merger equals its original employment share in the sum of the merged entities multiplied by the merged firm’s current employment level. The details of the imputation procedure are provided in the Appendix.

Young firms can be involved in a second or even more transfers in the first five years after entry. The probability of a second transfer is much higher (9.2%) than the unconditional probability of a first transfer (3.5%). Firms that have already relocated activities to another business number are much more likely to use this possibility again to optimize their legal or organizational structure. To err on the side of caution and not overestimate employment levels, we keep the employment of the original de novo entrant at the imputed level from the first transfer even after a second transfer. This assumption is likely to underestimate their growth rate as evidence from the first transfers suggests that employment on average increases after a transfer.

[25] The same figure for all entrants, i.e. not limited to de novo entrants, shows much higher employment in disappearing firm IDs, again concentrated in the two largest size categories.

Correctly classifying transfers as continuing firms lowers exit rates, especially in larger size classes. Imputing employment of firms involved in a transfer has only minor effects on the average growth rates of survivors. Firms involved in a transfer show on average similar imputed growth patterns as other firms in the same size class. Moreover, the group of surviving firms is much larger than the group of exiting firms, hence the lower impact on the average growth rate than on the average exit rate.

As mentioned before, firms with continued ID may be involved in a transfer as well, for example when they split off part of their activities to a new entity or take over another firm. Their employment is shown in the grey areas in Figure 2 and is disproportionately concentrated in larger size classes. Because employment is not lost or created but merely reallocated to another firm ID, it should not be considered job creation or destruction. In the case of split-offs, job destruction is overestimated; in the case of take-overs, job creation is. True employment growth of these firms can be biased in either direction and we will see below that the effect of such transfers on firm-level growth rates and on the growth-size relationship is minor.[26]

4 Empirical model

We characterize survival and growth patterns by size and age for young firms using the employment history of de novo entrants up to the moment of exit, both as identified previously. We follow Dunne, Roberts and Samuelson (1989) and decompose the mean growth rate of a class of firms into the growth rate of survivors, weighted by the probability of survival, minus the probability of exit. Three dependent variables are used: an exit indicator, employment growth conditional on survival, and a summary measure, the growth rate of all firms. We provide results for de alio entrants, and for de novo entrants without controlling for transfers in the Appendix. They illustrate how ignoring changes in firm identifiers and working directly with the administrative data hides the distinct growth and exit patterns of young firms.

Employment is measured as the number of employees registered on June 30. The set of entrants in year *t* includes all firms that started as an employer since July 1 of year *t*-1 and survive until June 30 of year *t*. It conditions on surviving a first selection process of unequal length, from a firm's establishment, the unknown point in time of age 0, to the first recorded instance of positive employment, denoted as age 1.[27] Firms with market success in their first year may infer above average efficiency and choose to expand before we observe them for the

[26] There are several reasons why firms might be involved in a transfer. Successful young firms are for example more likely to be taken over. Rapidly growing firms also have an incentive to split-up activities into smaller units to remain below the size threshold to avoid some legal obligations. In Belgium, small firms do not need to file full annual accounts or install a works council (fewer than 100 employees, turnover below 7.3m EUR, and balance sheet total below 3.65m EUR). See Garicano et al. (2013) for an illustration for French firms.

[27] The first five years after entry as an employer will be referred to as age 2 to 6.

first time. It rationalizes why firms have different sizes already at age 1 and why firms of different sizes exhibit different survival and growth rates already between ages 1 and 2.[28]

Exiters in observation period *t-1* to *t* are firms for which *t-1* is the last year of positive employment. Firms that change ID number, merge or split up, are not considered as exits. Their growth path following a transfer is based on imputed employment. The years between entry and exit, firms are denoted as survivors. Some survivors have zero employment in t-1 or t ('dormant' firms). They are treated as outliers and omitted from the regressions in the periods concerned.

Following Davis, Haltiwanger and Schuh (1996a), we calculate firm-level growth rates as discrete-time growth rates with average size in the denominator. Denoting employment of firm i in year t as E_{it}, its growth rate over the preceding year equals $g_{it} = (E_{it} - E_{it-1})/\bar{E}_{it}$, with $\bar{E}_{it} = (E_{it} + E_{it-1})/2$. These growth rates range from -2 for exits to +2 for entrants, show job creation and destruction symmetrically and are bounded away from infinity. Regressions use employment weights such that the estimates are readily interpreted as aggregate employment changes for a class of firms. Specifically, the mean estimated growth rates represent net employment growth for a given age-size class of firms, and the exit rates represent job destruction rates.

At each age, firms are grouped into five size classes: firms with one employee, 2 to 4 employees, 5 to 9 employees, 10 to 19 employees, and 20 or more employees.[29] Exiters are assigned to the size class of employment in their last year. For survivors, average employment in year t-1 and t is used to mitigate statistical side-effects from a conventional base-year classification. In particular, random fluctuations in size due to measurement error or transitory employment variation lead to regression-to-the-mean effects and induce a negative bias in the relation between size and growth (Friedman, 1992; Davis, Haltiwanger and Schuh, 1996b). Moreover, base-year size classes bound firm employment in the subset of survivors from below by one. Hence, the lower tail of possible rates of decline is truncated for smaller firms, e.g. one-employee firms that survive cannot have negative growth rates. It again induces an inverse relation between growth and size for survivors, even if employment fluctuations are independent of size (Baldwin and Picot, 1995).

To document patterns of firm dynamics, we regress firm-level employment growth or an exit dummy on age and size classes using a saturated dummy regression model.[30] It includes

---

[28] Firms may even have been in operation without employees for more than one year before age 1, either as a legal company or as self-employed. The decision to engage employees already reflects prior information about success.

[29] More detailed size classes above 20 employees are not needed because few de novo entrants reach this size within their first 5 years of operations. Due to averaging and employment imputation for transfers, employment is a continuous variable. Firms are classified into the following size intervals:]0,2[, [2,5[, [5,9[, [10,20[, [20,∝[. For example, the smallest size class]0,2[in period *t*-1 to *t* includes: exiters with one employee in *t*; stable survivors with one employee in *t*-1 and *t*; growing survivors from 1 to 2 employees; declining survivors from 2 to 1 employee.

[30] Regressions on de novo entrants include 365,000 firm-year observations; those for de alio entrants 44,000.

separate indicators for all possible values taken by the discrete explanatory variables age and size and their interactions. As Angrist and Pischke (2009) emphasize, saturated regression models fit the conditional expectation function perfectly, regardless of the distribution of the dependent variable. As mentioned before, using employment weighting makes the estimates represent net employment creation or destruction for each class of firms.

For each of the two dependent variables, $y_{it} = \{g_{it}, e_{it}\}$, employment growth and an exit dummy, the following regression model is estimated:

$$y_{it} = \sum_{j=1}^{5} \sum_{k=1}^{5} \left(\alpha_{jk} + \beta_{jk}^{d} D_{it}^{d}]\right) 1[age_{it} = j]\, 1[size_{it} = k] + \sum\nolimits_{d} \gamma_d D_{it}^{d} + \gamma_t + \varepsilon_{it}$$

where the dummy variable $1[age_{it} = j]$ takes a value of one if the age of firm i in year t equals j and similarly for the size category dummies. The industry dummies D_{it}^{d} enter both additively, to control for business cycle effects, and interacted with the full 5x5 set of age-size interactions. As we enforce $\sum_d \beta_{jk}^{d} = 0$, the average effect of age and size on growth and exit is captured by the uninteracted α coefficients, while the β^{d} coefficients allow for industry heterogeneity.

5 Results

5.1 Exit probability

Figure 3 plots the age-size coefficients for the exit regression which represent job destruction rates for each age-size class of de novo entrants. The coefficient estimates with standard errors are reported in Table A.3 in the Appendix (also for the following figures).

Many previous studies have found an inverse relationship between the exit rate and age, even conditional on size, and the results in the left panel confirm this pattern. Within every size class, failure rates decline almost linearly as firms age. Exit rates are especially high, at approximately 20 percent, in the first full year of existence. Five years after entry, they approximately halved, but they are still significantly higher than for incumbents, i.e. firms older than five years.

— Figure 3 approximately here —

The same coefficient estimates are plotted differently, in the right panel of Figure 3, to illustrate an inverse relationship between the exit rate and firm size within each age group. Exit rates decline with size and are especially high for one-employee firms. The pattern holds remarkably well for each age group and is even true for incumbents (the dashed line). The distances between the different lines are almost constant, consistent with a linearly declining exit rate with firm age for each size category.

The results are consistent with a model of passive learning which assumes that firms learn about their own efficiency level by operating in the market (Jovanovic, 1982). Profitable firms infer high efficiency and choose to continue and expand, while inefficient firms decide to exit. Over time only the more efficient firms from each entry cohort remain and exit rates decline. Within each age cohort, larger firms are those that have received favorable cost information in previous periods, and subsequent information gradually becomes less informative and less likely to induce exit. In contrast, firms that have grown more slowly are exactly those with worse information initially and they have lower value of staying in the industry. Another piece of negative information is more likely to induce them to exit. The results suggest that this process unfolds rather quickly in the very first few years after entry.

A remarkable feature of the patterns in Figure 3 is their regularity. The sample period spans the 2008-2009 recession which in Belgium lead to a 4 percent decline in real GDP, 12 percent fewer de novo entrants, and a one quarter increase in unemployment. Still, the exit rates decline monotonically with both age and size over the entire ranges and without a single exception. This regularity depends crucially on properly identifying true economic entry and exit, i.e. limiting the sample to de novo entrants and not classifying transfers as exits.

Results in Figure A.2 in the Appendix illustrate that the patterns are distorted if transfers are not filtered out of the exit statistics. The difference is not very large, but especially for larger size classes transfers induce an upward bias in exit probabilities. For firms of ages 2, 4 or 6, the exit probability in the largest size category is higher than for firms with 5 to 9 employees. Administrative exits for the largest firms are much more likely to involve a transfer of some of the work force, without discontinuation of activities.

In the second panel of Figure A.2, we also show the exit patterns for de alio entrants which enter with a new business number but are continuations of existing firms. Irrespective of firm age, the exit-size pattern is very much like that of incumbents. Within each size class, exit rates are much lower than for de novo entrants and they show little, or sometimes random, variation by age. Especially in larger size classes, where de alio entrants are overrepresented, mixing the two groups of firms would bias exit probabilities and hide the strong inverse relationships between exit and age that we documented for de novo entrants.

The narrow focus on de novo entrants and real exits leads to results that are more regular, but not altogether different from a straightforward analysis of administrative data. The reason is that de alio entrants and exits through transfers are concentrated in the same size classes. Ignoring changes in firm identifiers due to restructuring introduces two opposing biases that balance each other. Hence, exit patterns are relatively robust to misclassifications. It could explain the strong similarities found in many empirical studies on the relation between exit rates, size and age. As we show next, this is not the case for growth rates.

5.2 Growth of survivors

Figure 4 plots the coefficients from the employment growth regression of de novo entrants that survive between *t*-1 and *t*. They represent net employment growth rates of survivors within a given age-size class of firms.

— Figure 4 approximately here —

The results again show a negative relation with firm age conditional on size. In the early years after entry, surviving young firms of all sizes exhibit high growth rates, but these decline rapidly with age. The decrease in growth rates with age is again monotonic within every size class and the different lines in the left panel all shift down without crossing. In contrast with the exit probabilities, the decline is now convex as growth rates converge to a constant steady state. The strongest reduction occurs in the first few years after entry. A second difference is that the eventual growth rates for the different size classes converge to very similar, slightly negative growth rates. On average, growth rates of surviving young firms at age 6 are still positive and about 4 percentage points higher than for mature firms of the same size. For the smallest firms growth has basically stalled after five years while for larger firms growth will remain positive for a few years longer. It implies that the firm distribution will continue to shift to the right.

The inverse relation between growth rates and age of surviving firms is also consistent with the passive learning model (Jovanovic, 1982). It assumes that a firm typically enters small and as it learns its true costs it adjusts its size towards the profit maximizing level. Firms that discover they are more efficient, grow and survive, while the inefficient shrink. Firm sizes within an entry cohort will diverge to some extent, but the weakest firms will eventually decide to exit the industry. As time passes, additional information becomes less informative and firms converge to their optimal size with no further growth (on average).

Firm growth by size in the first years after entry

The different experience by firm size class that could be inferred from the ordering of the lines in the left panel is highlighted in the right panel of Figure 4. Among firms of the same age, growth rates are strongly increasing in firm size. The positive relationship is especially pronounced in the very first years after entry, and moves towards more proportional growth rates as a cohort matures. The dispersion in growth rates across size classes reduces gradually with firm age and become statistically insignificant by age 6. These results suggest that Gibrat’s law is strongly violated for surviving young firms of the same age.

As shown below, the positive size-growth relationship is entirely absent if de alio entrants are not adequately filtered out from the population of entrants. This is one of the main reasons why our results differ from earlier studies. For the population of de novo entrants, however, our results are robust across different subpopulations and to alternative size classifications.

The increasing growth rates by size of young firms in the very first years after entry is a striking result and challenges earlier research which finds an inverse relation with size (Evans, 1987a; Dunne, Roberts and Samuelson, 1989; Mata, 1994). The differences with these findings is explained by three elements. First, we focus on the very first years after entry, while already at age 5, growth rates move towards a more proportional distribution. Second, we include the smallest entrants even one-employee firms, while the increasing growth rates with size are concentrated among firms with fewer than 20 employees. Third, and most importantly, we focus on de novo entrants, excluding de alio entrants, which are continuations of older firms. Including the latter would strongly bias growth rates downward in larger size classes, making the positive growth-size relationship disappear entirely.

Figure A.3 in the Appendix successively applies different adjustments. It shows, in particular, the incumbent-like growth patterns for de alio entrants which bias growth rates down in large size classes, where they are overrepresented. They only grow faster than incumbents in the first recorded year and the positive growth-size relationship is barely perceptible for any age cohort, in line with the evidence for incumbents. The administratively recorded age of these firms seems unrelated to actual firm age. De alio entrants are already in a more advanced stage of the selection process with less need for size adjustments. Therefore, analyzing administrative entrants without accounting for firm restructuring would incorrectly indicate that the proportional or decreasing growth rates with size, as is found for mature firms, also applies to young firms in the early years after entry.

Many studies have found that smaller firms have higher growth rates and this also holds in our data if we pool across age groups. This pattern is consistent with a positive growth-size relation conditional on age because surviving young firms grow very strongly and they tend to be small. A few empirical studies have even found an inverse relation between growth and size holding age fixed. The difference with our results can be explained by their focus on large entrants (Mata, 1994), a coarse age classification (Dunne, Roberts and Samuelson, 1989), not controlling for the effects of regression to the mean (Evans, 1987a; Lotti, Santarelli and Vivarelli, 2001) and, most importantly, not controlling for entrants that emerge from ID changes or firm restructurings. Using U.S. census data with improved firm links, Haltiwanger, Jarmin and Miranda (2013) and Decker et al. (2013) also found a mild positive relation between growth and size for surviving firms. Our results complement them and show that a narrow focus on firms of the same exact age leads to a very pronounced positive relation, but only in the first years after entry.

The positive growth-size relation is consistent with the model of passive learning in Jovanovic (1982) if we take the entry date to lie some unknown period before the first time firms are observed with positive employment (always on June 30). Firms all enter with the same size, but sizes diverge as firms receive imperfect information about their own quality from their market operations. Firms receiving positive draws adjust the assessment of their own productivity using a Bayesian updating mechanism and expand. This adjustment is only gradual as firms still place some weight on prior information. If, as suggested by Mata, Portugal and Guimaraes (1995), adjustment costs prevent young firms to adjust their size instantaneously, current size is likely to be different from desired size. For expanding firms, growth rates will be higher for some time to incorporate all past information and overcome adjustment barriers, leading to the observed positive relation.

The estimated patterns reflect several aspects of this early adjustment process. Most de novo entrants indeed enter with the same, very low size: 94 percent have fewer than 5 employees and entrants with more than 50 employees are extremely rare. The observed variation is plausibly the result of small differences in entry dates. The different observed entry sizes, moreover, have a positive predictive power for survival and growth rates in the first years. We interpret this is a gradual adjustment process to past positive information. At each point in time, larger firms are the ones that received the most positive information about their capabilities and have chosen to expand. Their subsequent growth and survival rates to some extent still reflect adjustment to this information. As revisions of estimated efficiency become smaller over time, firms approach their optimal scale. We do observe growth rates to move towards a more proportional growth distribution as firms mature.

Robustness checks

We conduct several robustness checks that confirm the positive relationship between growth and size of de novo entrants – these are reported in the Appendix. Not correcting for transfers has only a very small impact on the empirical pattern (Figure A.3). Excluding de novo entrants that are at some point involved in a transfer of activities after entry (ID-change, merger, spin-off) does not change the pattern either. For these firms employment has been imputed, but omitting them even makes the positive relationship slightly more pronounced.

The estimated growth rates are averages over cohorts of firms that entered between 2003 and 2011. This period includes the severe 2008-09 recession which is only controlled for using additive time dummies. To verify whether post-entry growth patterns are sensitive to the business cycle, we estimated the model separately for firms entering in the period of high economic growth before 2008 and firms entering during the recession or in the period of slow recovery afterwards. In spite of the markedly different economic environment, the absolute growth rates as well as the patterns by size and age are remarkably similar in both periods (Figure A.4). The experience of de novo entrants seems rather insensitive to the business cycle.

It is well known that regression to the mean will spuriously induce a negative relationship if firm size is measured at the start of the period over which growth rates are calculated (Davidsson, Lindmark and Olofsson, 1998). We have followed the solution of Haltiwanger, Jarmin and Miranda (2013) and used the average size in years t and t-1 as a proxy for the size over the t-1 to t period. If firm size fluctuated around a stable long-run size, using the average size classification would yield unbiased results. However, in a sample with on average positive growth, it introduces a positive bias between growth and size for firms crossing size class borders. To verify how important this effect is, we use two alternative ways to classify firms by size that approximate the continuous growth-size relationship.

First, we use a dynamic size classification that allocates a firm's employment increase or loss to the size class in which the change occurs. The method is used by the U.S. Bureau of Labor Statistics to avoid base-year classification biases in the Business Employment Dynamics statistics (Butani et al., 2006). Firms are initially assigned to a size class based on their employment in t-1, E_{it-1}, but are re-assigned to a new class when the class threshold is crossed. The growth from E_{it-1} to the threshold is assigned to the initial class and the growth from the threshold to E_{it} is assigned to the next size class.[31] It approximates instantaneous re-assignment that would be feasible if size and growth were measured in continuous time.

Second, our use of average firm size to approximate the continuous growth-size relationship can be motivated similarly as the use of average wage shares in a Solow-residual,

[31] The last employee that moves the firm across a threshold is allocated to the smallest of the two size classes.

i.e. as a discrete approximation to the continuous Divisia index of productivity growth. As illustrated by Caves, Christensen and Diewert (1982), the same index can also be interpreted as a geometric mean of two productivity comparisons, each using the technology of a different base year (t-1 or t). In our application, we can calculate both the growth rate and the firm size class first using employment at t-1 as base, i.e. $g_{it} = (E_{it} - E_{it-1})/E_{it-1}$ and $1[E_{it-1} = j]$, and a second time using E_{it} as base for both variables. We then run the growth regression giving the measures using E_{it-1} or E_{it} as base each a weight of one half.

The results using both robustness checks are shown in the two panels of Figure 5. For the dynamic size classification we have chosen slightly different size-class thresholds to allow symmetric and equal ranges of growth rates between -0.67 and +0.67 within each class. For both methods the positive relationship between employment growth and firm size conditional on age survives. The absolute growth rates for the youngest firms are lower, but the curve shifts down over the entire size distribution. For each age cohort, growth rates still increase monotonically with size without a single exception. The patterns look even more regular than in Figure 4. For the dynamic size classification, the positive relationship does become less pronounced already after a firm's fourth full year of operation.

— Figure 5 approximately here —

Finally, we verify whether the positive relationship between growth and size holds in all sectors. The empirical specification included six additive industry dummies as well as interactions with the age-size variables. The coefficients on the interactions were normalized to sum to zero, such that the uninteracted age-size coefficients represented average growth rates for a class of firms in the total economy, giving each industry equal weight. In Figure A.5, we show separate results for six industries: Manufacturing, Construction, Trade, Accommodation & Food Services, Business Services, and Mixed Household & Business Services.

Growth rates of surviving de novo entrants show broadly the same patterns as in Figures 4 and 5 in all industries. They are high in the first year, but decrease quickly with age within each size class. Only in Accommodation & Food Services, where average firm size is small, there is little room for size diversification after entry resulting in little difference between firms from age 3 onwards. In all other industries, we see the same positive relation between growth and size conditional on age, while for older cohorts the pattern moves to a more proportional distribution. The increasing relationship tends to be more pronounced in service sectors, where entry costs are often lower. Firms can easily enter with a very small size and gradually adjust to an optimal scale. Consistent with a higher minimum efficient scale in manufacturing, we find higher average size at entry and a negative growth-size relation for size classes above 10 employees. In the least cyclical sector, Mixed Household & Business Services, growth rates are monotonically increasing with size at every single age.[32]

[32] The differences between industries seem unrelated to aggregate industry growth or average firm size: Mixed Services and Manufacturing have been shrinking over the sample period, while Construction and

5.3 Growth of all entrants – Net employment creation

Entrants contribute positively to aggregate employment by the jobs they create in the year of entry and subsequently if employment growth for survivors outweighs job destruction by exiters. We first illustrate the contribution of de novo entrants of specific age-size classes using the same specification as before, but estimating the regression for all firms, not conditional on survival. We then illustrate the evolution over time of aggregate job creation of a representative cohort of entrants.

Figure 6 shows the combined effect of the survival and conditional growth patterns that we documented in Figures 3 and 4. On the left, the declining exit rate with age (increasing survival rate) is balanced by the declining growth rate as surviving entrants mature. It results in an almost flat pattern in most size classes. Only for one-employee firms does the rapidly declining exit rate dominate and is the net employment creation profile increasing with age. The probability of survival and net growth rates are very low, however, such that even at age 6 the net contribution to employment creation of these small firms is still negative, only less negative than at the start.

For the larger size classes, the declining growth rate dominates in the first two years—Figure 4 already showed a halving of firm-level growth from age 2 to age 3. From age 4 onwards, net employment creation by size class remains approximately constant as reduced exit is compensated by lower growth rates for survivors. While the growth rates for firms at age 6 are still statistically significant from incumbents in the same size classes, this is due to very small standard errors (see Table A.3) as the absolute differences are relatively small.

The ranking of the different size classes is interesting. In the steady state, net employment contribution of each group is negative, but especially so for smaller size classes. For the largest firms, de novo entrants that expanded to at least 20 employees by age 6, net employment growth is approximately zero. For all other size classes it is significantly negative: around –8 percent for firms with a workforce of 5–9 or 10–19 employees, –13 percent for firms with 2–4 employees, and –27 percent for single-employee firms.

— Figure 6 approximately here —

These same estimates are again presented as a function of firm size in the right panel. The declining exit rates by size now reinforce the increasing growth rate for survivors. It leads to an even stronger positive growth-size relationship. Gibrat's Law is violated even more starkly if one does not condition on firm survival. Because growth rates fall with age while survival rates increase and these relationships are equally strong over the entire size spectrum, the net employment curves for the different age classes are almost on top of one another. Only for the youngest firms is the net growth-size relationship steeper, reflecting the more pronounced growth-size relationship conditional on survival.[33]

Business services were expanding; in both Mixed Services and Manufacturing the average firms size is large, while it is small in Construction and Trade, without systematic differences in the growth patterns.

[33] For incumbents, the relationship is slightly less steep, again due to the pattern for conditional growth rates.

Net growth rates are positive for some time as firms mature, and this holds for a longer period in larger size classes. Even for firms with 5 to 9 employees, net growth rates already turn negative by age 3. This pattern has implications for the evolution of net employment creation of the entire entry cohort. If the share of small firms, which is extremely large at entry, remains sufficiently large, employment created at entry will fall continuously as cohorts age. If in contrast, the size distribution moves sufficiently rapidly to the right, the positive net growth rates of larger firms could yield a net increase of aggregate employment for some years after entry.

The left panel of Figure A.6 in the Appendix shows the evolution of the firm size distribution as de novo entrants mature. Similar to Cabral and Mata (2003), the strongly right-skewed distribution at entry gets a fatter right tail as the cohorts matures. While small firms dominate initially, by age 6 the distribution has almost fully converged to that of incumbents. The pronounced shift makes the positive growth-size relation for survivors, as documented in Figure 4, disappear if one pools firms of different ages. The greater importance of older firms in the larger size classes even leads to a negative growth-size relationship without controlling for age (right panel of Figure A.6), a pattern found in many earlier empirical studies.

Figure 7 shows the employment evolution of a representative cohort of de novo entrants. It combines the effect of the age-size specific exit and growth rates and the shifting size composition as firms mature. Similarly as Boeri and Cramer (1992), we find that a cohort's employment remains stable for one year, but decreases monotonically with age thereafter. Five years after entry, employment already falls to 88 percent of its initial value. The firm size distribution does not shift to larger size classes sufficiently quickly to compensate for the exit of some firms and the slowing growth rates of survivors.

— Figure 7 approximately here —

Initial size at entry has some predictive power for subsequent employment growth. For the large subset of de novo entrants that start with fewer than 5 employees (94 percent of firms) total employment declines even more quickly. Five years after entry, it stands at 86 percent of its initial level. In contrast, employment by larger entrants increases in the first year and decreases at a much slower pace afterwards. Five years after entry, employment for this group of entrants is still 95 percent of the initial level. It is consistent with the interpretation of initial size that we suggested before. Employer size at entry is the result of an initial selection process even before we first observe the firm. Larger entrants are the more efficient survivors of this initial process. They will have lower post-entry exit rates and their growth rates do not decline as quickly. As a result, they are better able to maintain the employment level at their recorded entry time than smaller entrants.

The same evolution can be calculated on administrative data directly, without distinguishing de novo from de alio entrants or correcting for transfers. The dashed line in the left panel shows that total employment for such entry cohorts declines even more quickly and more pronouncedly. The classification errors induce three biases. Including incumbent-like de alio entrants lowers employment losses due to exit, but also the employment growth of surviving firms. Misclassifying transfers as exits overestimates job losses. The latter two effects dominate and the employment reduction as the cohort matures is overestimated.

Calculating separate employment evolutions by size of entry also on the administrative data reveals two interesting patterns. First, compared to de novo entrants, the effect of initial size reverses, now showing a more pronounced employment decline for larger entrants. It incorrectly supports the perception that initial size is negatively related to post-entry performance. Second, both entry definitions leads to very similar employment evolutions for small entrants, but not at all for larger entrants. Misclassifying entry and exit are both much more likely for larger firms. The downward bias for administrative entrants is entirely concentrated with the group of larger entrants.

6 Conclusion

We have shown that a reliable assessment of the importance of job creation by new entrants hinges crucially on clean identification of entry and exit. The overall evolution of job creation combines the impact of each of the items discussed in this paper: the initial size distribution of entrants, the patterns of firm exit and growth conditional on survival by age and size category.

Focusing on a well-defined set of *de novo* entrants clearly shows that initial job creation of truly new firms is quite modest in Belgium, each year accounting for only 1.5 percent of aggregate employment. As exit rates are very high initially, especially for small firms that make up the bulk of new entrants, maintaining the initial employment level is only possible during the first year when employment growth rates for survivors are still very high. As firms mature, average growth rates decline and the effects of firm exit quickly start to dominate.

Two patterns help sustain initial employment levels. First, as surviving firms gradually shift into larger size classes, the positive relationship between firm growth and size sustains employment growth for survivors. Second, properly identifying transfers of activities to newly formed enterprises and not classifying such events as economic exits lowers measured job destruction rates. Both adjustments are more important for larger firms. As a result, the group of entrants that start with a minimum of 5 employees are able to sustain their cohort's initial employment level for four years after entry.

These results are remarkably regular, even over the very turbulent sample period. In particular, the increasing relationship between growth and firm size in the first years after entry holds without a single exception in all size categories. It even holds in all robustness checks using various alternative measures and for firms entering in the high or the low growth period. The patterns for survival rates, increasing in size conditional on age and increasing in age conditional on size, are equally robust, but previous studies already highlighted these patterns. The findings for de novo entrants are in sharp contrast with those for de alio entrants which behave almost like incumbents. Therefore, distinguishing the two types of entrants is especially important for correctly identifying the positive growth-size relationship which only holds in the firms first few years of operations.

The increasing growth-size relationship contrasts with the finding of most earlier literature. It can be explained by a redirection of the focus of analysis from larger to smaller entrants. On the one hand, we identify a majority of entrants in larger size classes as continuations of existing firms. On the other hand, we include a mass of small de novo entrants which have been neglected in most earlier literature. These smaller entrants predominantly stay small or exit at early age.

In terms of policy conclusions, a few patterns are worth emphasizing. The popular claim that "small firms are the engines of job creation" is not false as an unconditional empirical relationship, but it is highly misleading. Firm growth is strongly increasing in firm size especially in the first years after entry. The effect of higher survival rates by firm size further strengthens this positive relationship. The absolute level of job creation by entrants is also much smaller than is often appreciated. Half of the jobs in the group of administrative entrants is accounted for by de alio entrants, firms that are founded out of existing firms. At the same time, de novo entrants that employ at least 5 workers contain many promising firms. In particular, they have very high organic growth rates. Therefore, policy measures aimed at supporting new firms should not have upper limits on firms size.

Appendix – Imputation of employment after transfers

First transfer

Employment of de novo entrants involved in a first transfer is imputed as follows. Let s be the initial ID number of the de novo entrant and $i = 1,2,\ldots n$ the ID number(s) of the firms it is linked to in period $t\text{-}1$ to t. Let $E_{s,t}$ and $E_{i,t}$ be observed administrative employment in year t of s and i respectively, and $F_{s,t+n}$ imputed employment of s in year $t+n$ for $n \geq 0$.

ID-change: a transfer is defined as an ID-change if the ID-number s of the de novo entrant exits and is linked to the ID-number i of a firm that enters in t. Imputed employment of s in the years after the transfer is given directly by the employment of i. It is defined as:

$$F_{s,t+n} = E_{i,t+n} \qquad \text{for } n \geq 0 \tag{A1}$$

Split-up: a transfer is defined as a split-up if the ID-number s of the de novo entrant exits and is linked to the ID-numbers i of more than one firms that enter in t; or, if the ID-number s of the de novo entrant continues and is linked to the ID-number i of at least one firm that enters in t. Imputed employment of s in the years after the transfer is equal to the sum of employment of s and i. It is defined as:

$$F_{s,t+n} = E_{s,t+n} + \Sigma_i E_{i,t+n} \qquad \text{for } n \geq 0 \tag{A2}$$

Merger: a transfer is defined as merger if the ID-number s of the de novo entrant exits and is linked to the ID number i of an incumbent (s is absorbed by i); or, if the ID-number s of the de novo entrant continues and is linked to the ID number i of at least one firm that exits (s absorbs i); or, if the ID-number s of the de novo entrant and the ID-number i of at least one other firm, which all exit, are linked to one and the same ID number j of a firm that enters in t (s and i are merged into a new entity j). Let k be the ID number of the merged entity in each of the three cases, then imputed employment of s in the years after the transfer is equal to its employment share in the sum of the separate entities $s + i$ just before the merger, multiplied by the number of jobs of the new entity k. It is defined as:

$$F_{s,t+n} = [E_{s,t-1} \,/\, (E_{s,t-1} + \Sigma_i E_{i,t-1})\,] * E_{k,t+n} \qquad \text{for } n \geq 0 \tag{A3}$$

Imputed employment $F_{s,t+n}$ in all types of transfers can be summarized with a general expression which is also used to treat multiple transfers between firms in the same period. We use the notion 'total event' to denote the combination of all firms that are interlinked in one period $t\text{-}1$ to t. Consider all firm ID numbers $i = 1,2,\ldots n$ that are involved in the same 'total event' as the de novo entrant s in period $t\text{-}1$ to t. Imputed employment of s in the years after the transfer is equal to its employment share in the sum of the separate entities $s + i$ just before the merger, multiplied by the sum of employment of $s + i$. It is defined as:

$$F_{s,t+n} = [E_{s,t-1} \,/\, (E_{s,t-1} + \Sigma_i E_{i,t-1})\,] * (E_{s,t+n} + \Sigma_i E_{i,t+n}) \quad \text{for } n \geq 0 \tag{A4}$$

Second transfer

If the de novo entrant, or one of the other firms involved in the first transfer is linked to another business number in a subsequent period, imputed employment will in turn exhibit a spurious shock. Such events are denoted as second transfers.

Reconstruction of employment after this event is however not performed, since it would require to trace the list of all firms involved in a second ‘superevent’. Instead, employment of de novo entrants after a second transfer is kept constant from the beginning of the second transfer onwards. Given the strong growth rates of these firms prior to the second transfer, this estimation should be considered as rather cautious.

References

Albaek, K. and B.E. Sorensen (1998). “Worker flows and job flows in Danish manufacturing, 1980-91.” *Economic Journal*, 108, 1750–1771.

Angrist J.D. and J.S. Pischke (2009). *Mostly Harmless Econometrics. An Empiricist’s Companion*. Princeton: Princeton University Press.

Baldwin, J.R., R. Dupuy, and W. Penner (1992), “Development of longitudinal panel data from business registers: the Canadian Experience.” *Statistical Journal of the United Nations*, 9, 289–303.

Baldwin, J.R., D. Beckstead, and A. Girard (2002). “The importance of entry to Canadian Manufacturing.” OECD STI Working Paper No. 189. Paris: OECD.

Baldwin, J.R. and G. Picot (1995). “Employment generation by small producers in the Canadian manufacturing sector.” *Small Business Economics*, 7:4, 317–331.

Bartelsman, E., J.C. Haltiwanger, and S. Scarpetta (2009). “Measuring and analyzing cross-country differences in firm dynamics.” In *Producer Dynamics: New Evidence from Micro Data* (eds. Dunne, Jensen, and Roberts), University of Chicago Press.

Benedetto, G., J.C. Haltiwanger, J. Lane, and K. McKinney (2007). “Using worker flows to measure firm dynamics.” *Journal of Business and Economic Statistics*, 25:3, 299–313.

Boeri, T. and U. Cramer (1992). “Employment growth, incumbents and entrants: Evidence from Germany.” *International Journal of Industrial Organization*, 10:4, 545–565.

Butani, S.J., R.L. Clayton, V. Kapani, J. Spletzer, D.M. Talan, and G.S. Werking Jr. (2006). “Business employment dynamics: Tabulations by Employer Size.” *Monthly Labor Review*, 2:1, 3–22.

Cabral L.M.B. and J. Mata (2003). “On the evolution of the firm size distribution: Facts and theory.” *American Economic Review*, 93:4, 1075–1090.

Campbell, J.R. (1998). “Entry, exit, embodied technology, and business cycles.” *Review of Economic Dynamics*, 1, 371–408.

Caves, D.W., L.R. Christensen, and W.E. Diewert (1982). “The economic theory of index numbers and the measurement of input, output, and productivity.” *Econometrica*, 50:6, 1393–1414.

Davidsson, P., L. Lindmark, and C. Olofsson (1998). “The extent of overestimation of small firm job creation – An empirical examination of the regression bias.” *Small Business Economics*, 11:1, 87–100.

Davis, S.J., J.C. Haltiwanger and S. Schuh (1996a). *Job creation and destruction*. Cambridge: MIT Press.

Davis, S.J., J.C. Haltiwanger and S. Schuh (1996b). “Small business and job creation: Dissecting the myth and reassessing the facts.” *Small Business Economics*, 8:4, 297–315.

Decker, R., J. Haltiwanger, R. Jarmin and J. Miranda. “The role of entrepreneurship in US job creation and economic dynamism.” *Journal of Economic Perspectives*, 28:3, 3–24.

Dunne, T., M.J. Roberts, and L. Samuelson (1989). “The growth and failure of U.S. manufacturing plants.” *Quarterly Journal of Economics*, 104:4, 671–698.

Eriksson, T. and J.M. Kuhn (2006). “Firm spin-offs in Denmark 1981–2000. Patterns of entry and exit.” *International Journal of Industrial Organization*, 24, 1021–1040.

Eurostat (2003). *Business Register: Recommendations Manual*. Luxembourg: Office for Official Publications of the European Communities. (Available online at http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-BG-03-001-__-N-EN.pdf

Eurostat-OECD (2007). *Manual on Business Demography Statistics*. (Available online at http://www.oecd.org/dataoecd/8/8/39974460.pdf).

Evans, D.S. (1987a). "The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries." *Journal of Industrial Economics*, 35:4, 567–581.

Evans, D.S. (1987b). "Tests of alternative theories of firm growth." *Journal of Political Economy*, 95:4, 657–674.

Friedman, M. (1992). "Do old fallacies ever die?" *Journal of Economic Literature*, 30:4, 2129–2132.

Garicano, L., C. Lelarge, and J. Van Reenen (2013). "Firm size distortions and the productivity distribution: Evidence from France." NBER Working Paper No. 18841.

Geurts, K. and P. Vets (2013). "Employee flows to study firm and employment dynamics." DynaM Working Paper, KU Leuven.

Haltiwanger, J., R.S. Jarmin, and J. Miranda (2013). "Who creates jobs? Small versus large versus young." *Review of Economics and Statistics*, 95:2, 347–361.

Hethey T. and J. Schmieder (2013). "Does the use of worker flows improve the analysis of establishment turnover? Evidence from German administrative data." NBER Working Paper No. 19730.

Hopenhayn, H. (1992). "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica*, 60:5, 1127–1150.

Hsieh, C.-T. and P. Klenow (2014). "The life cycle of plants in India and Mexico." *Quarterly Journal of Economics* (forthcoming).

Jarmin, R.S. and J. Miranda (2002). "The Longitudinal Business Database." Center for Economic Studies Discussion Paper CES-WP-02-17.

Jovanovic, B. (1982). "Selection and the evolution of industry." *Econometrica*, 50:3, 649–470.

Lotti, F., E. Santarelli, and M. Vivarelli (2003). "Does Gibrat's Law hold among young, small firms?" *Journal of Evolutionary Economics*, 13:3, 213-235.

Mata, J. (1994). "Growth during infancy." *Small Business Economics*, 6:1, 27–39.

Mata, J. and P. Portugal (2004). "Patterns of entry, post-entry growth and survival." *Small Business Economics*, 22:3-4, 283–298.

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

Mead, D.C. (1994). "The contribution of small enterprises to employment growth in Southern and Eastern Africa." *World Development*, 22:12, 1881–1894.

Mikkelson, G., L.I. Unger, and D. LeBel (2006). "Identifying and accounting for mergers and acquisitions in measuring employment." BLS Statistical Survey Papers. Washington.

Muendler, M.A., J.E. Rauch, and O. Tocoian (2012). "Employee spinoffs and other entrants: Stylized facts from Brazil." *International Journal of Industrial Organization*, 30, 447–458.

OECD (2013). *Entrepreneurship at a Glance*. Paris: OECD Publishing.

Rollin, A.M. (2013). "Developing a longitudinal structure for the national accounts longitudinal microdata file (NALMF)." Proceedings of Statistics Canada Symposium 2013.

Spletzer, J.R. (2000). "The contribution of establishment births and deaths to employment growth." *Journal of Business and Economic Statistics*, 18:1, 113–126.

Sutton, J. (1997). "Gibrat's legacy." *Journal of Economic Literature*, 35:1, 40–59.

Table 1: Share of de alio entrants in all administrative entrants

		Total	Firm size class (employees)				
			1-4	5-9	10-19	20-49	50+
Total share of *de alio* entrants (by either method)		0.09	0.05	0.36	0.59	0.74	0.89
1. Identified by official method		0.06	0.05	0.12	0.16	0.21	0.37
2. Identified by employee flow method		0.05		0.32	0.57	0.72	0.88
by type of employee flow link							
2.a	Predecessor and entrant share at least 50% of employees (ID-change)	0.04		0.25	0.44	0.54	0.65
2.b	At least 75% of entrant's employees are transferred from predecessor (Split-up)	0.01		0.05	0.1	0.16	0.22
2.c	Other	0.00		0.01	0.03	0.03	0.02

Note: Average of annual shares over the 2004-2012 period. Share of the number of firms, not of employment.

Table 2. Summary statistics for different types of entrants

	Share of entrants		Entry rate (%)	Job creation rate (%)	Average size	Size relative to incumbents
	Firms	Empl.				
All administrative entrants	1	1	9.3	2.6	3.1	0.26
- *De alio* entrants	0.09	0.44	0.9	1.1	14.6	1.22
- *De novo* entrants (by either method)	0.91	0.56	8.4	1.5	1.9	0.16
- Identified by official method	0.94	0.82	8.7	2.1	2.7	0.23
- Identified by employee flow method	0.95	0.59	8.8	1.6	2.0	0.16

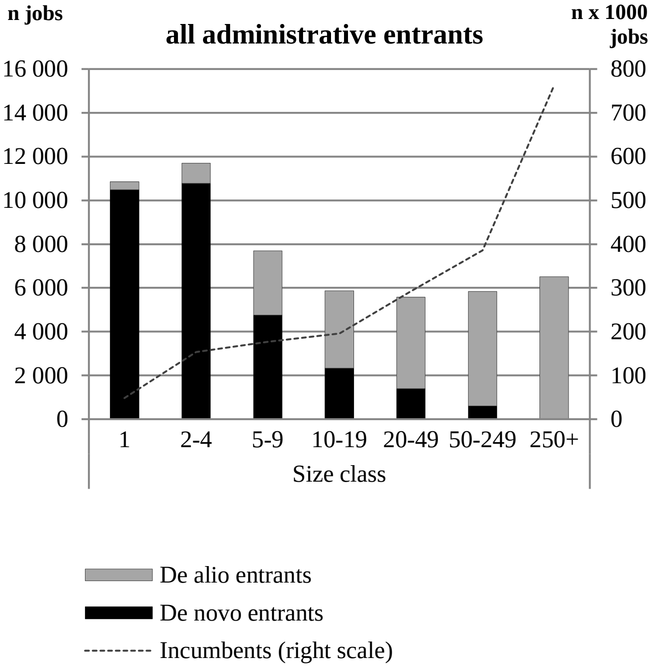
Note: Annual averages over the 2004-2011 period. Entry rate is % of active firms; Job creation rate is % of total employment; Average size is measured in employees; Size relative to incumbents is the ratio of mean firm employment of each group to that of incumbent firms.

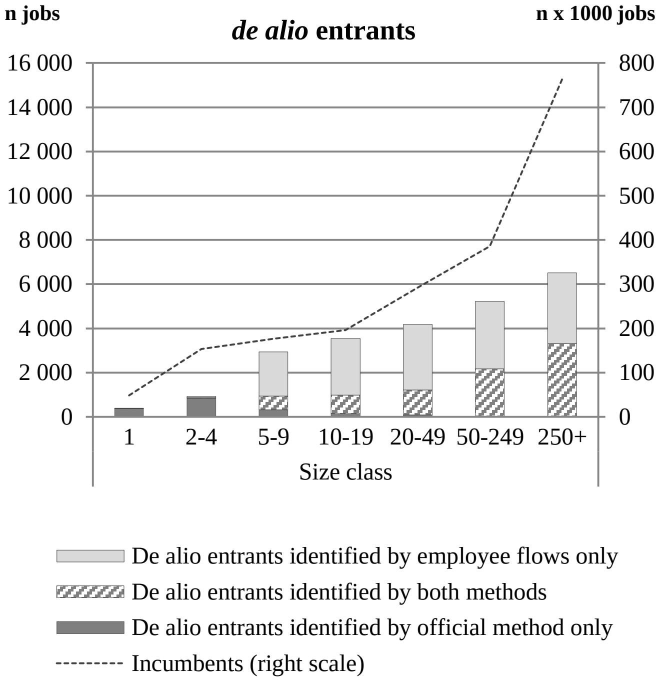
Table 3: Share of transfers in all administrative exits of de novo entrants 1 to 5 years old

		Firm size class before exit (employees)				
	Total	1-4	5-9	10-19	20-49	50+
Total share of *transfers*	0.04	0.02	0.17	0.29	0.34	0.41
1. Identified by official method	0.03	0.02	0.05	0.07	0.07	0.10
2. Identified by employee flow method	0.01		0.15	0.27	0.33	0.41
by type of employee flow link						
2.a One-to-one ID change	0.01		0.12	0.20	0.17	0.24
2.b Split into several business numbers	0.00		0.00	0.01	0.07	0.14
2.c Merged to another business number	0.00		0.03	0.06	0.09	0.03

Note: Average of annual shares over the 2004-2012 period. Share of the number of firms, not of employment.

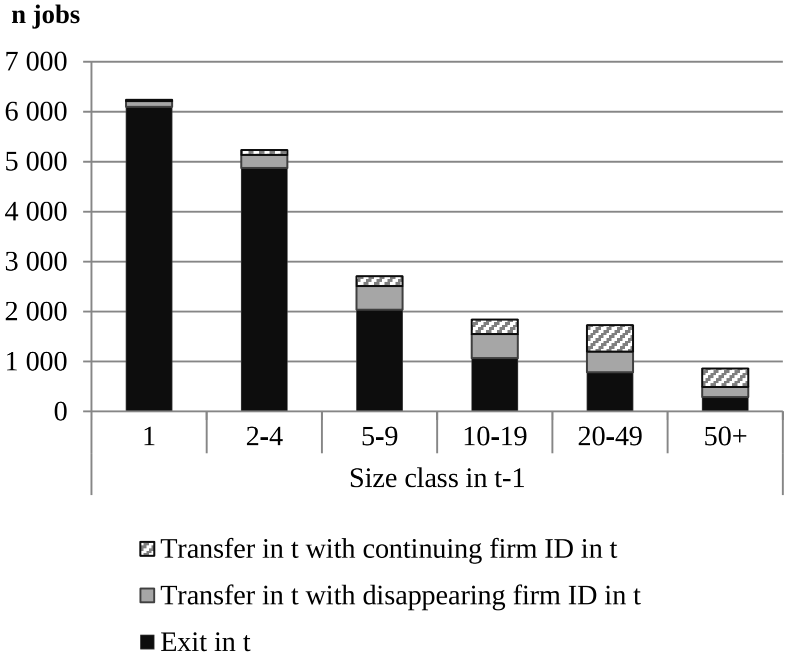
Figure 1: Employment distribution of de novo and de alio entrants





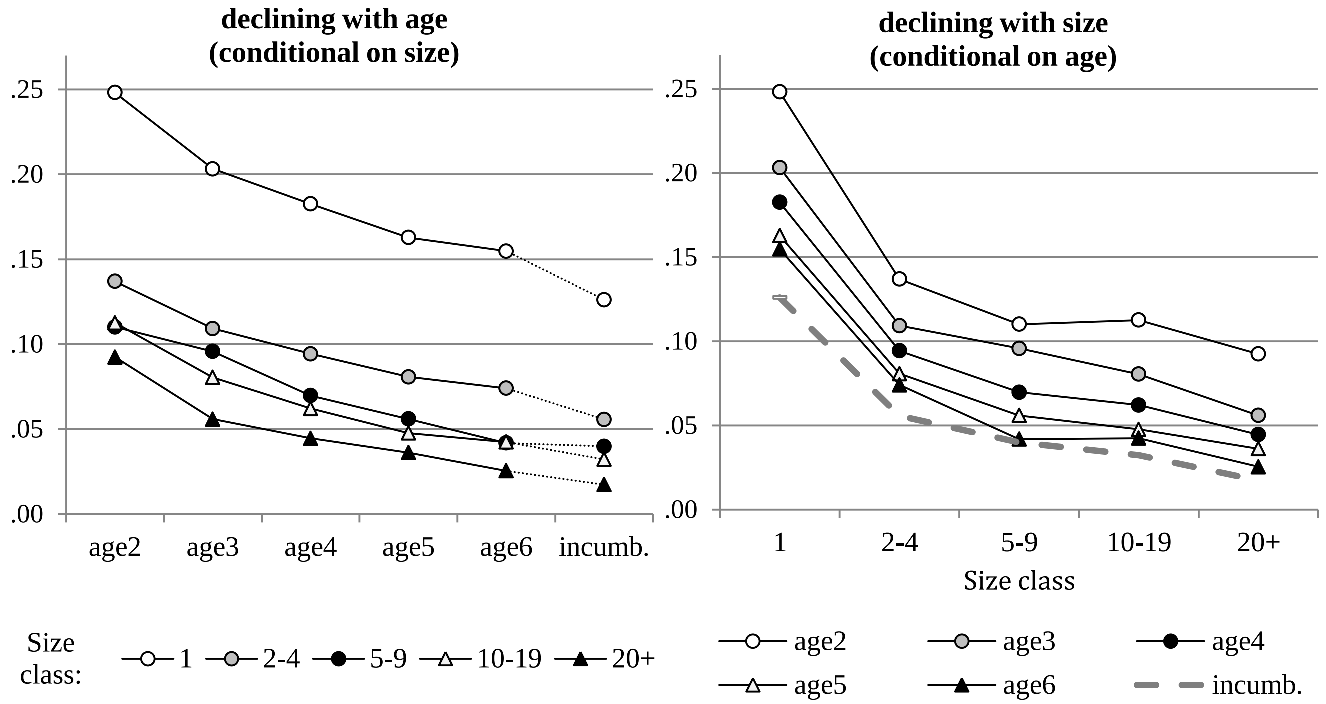
Note: Average of annual entry distributions over the 2004-2012 period.

Figure 2: Employment distribution for de novo entrants in year before exit or transfer



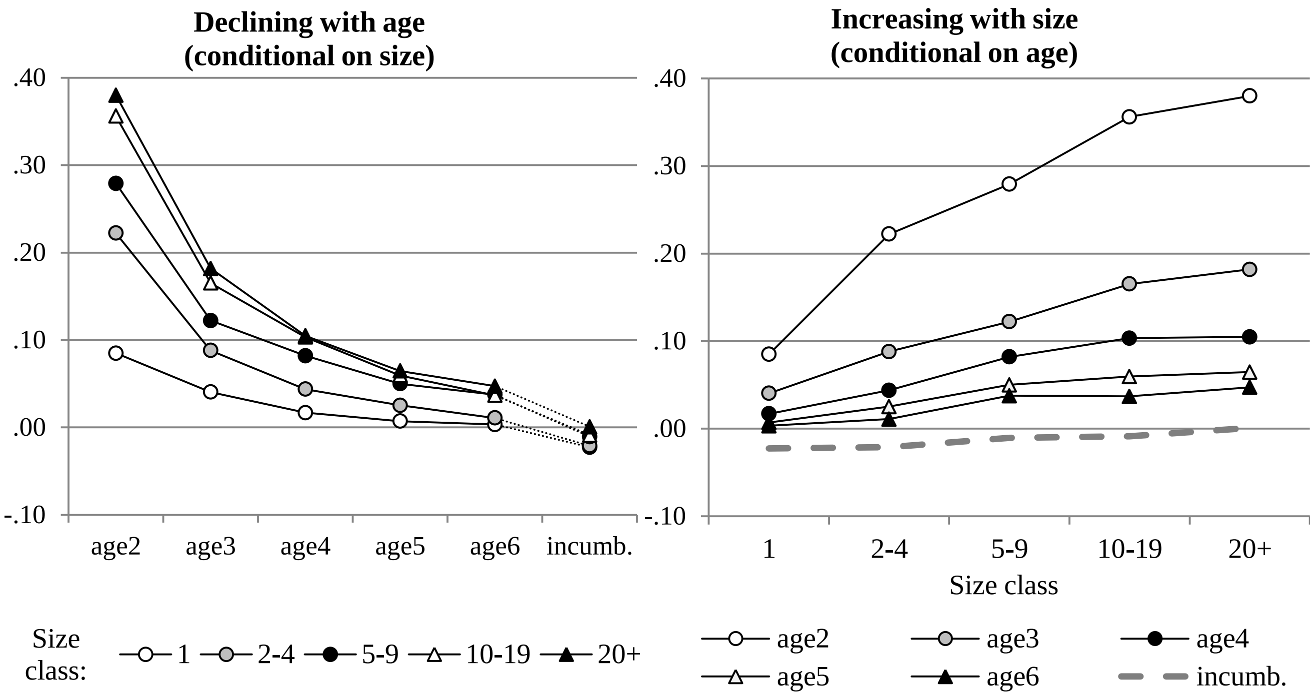
Note: Average of annual entry distributions over the 2004-2012 period.

Figure 3: Exit rates of de novo entrants



Note: Regression coefficients estimated over the full sample period (2004-2012) using employment weights. Both panels report the same coefficients. Incumbents are firms older than 6 years.

Figure 4: Growth rates of surviving de novo entrants

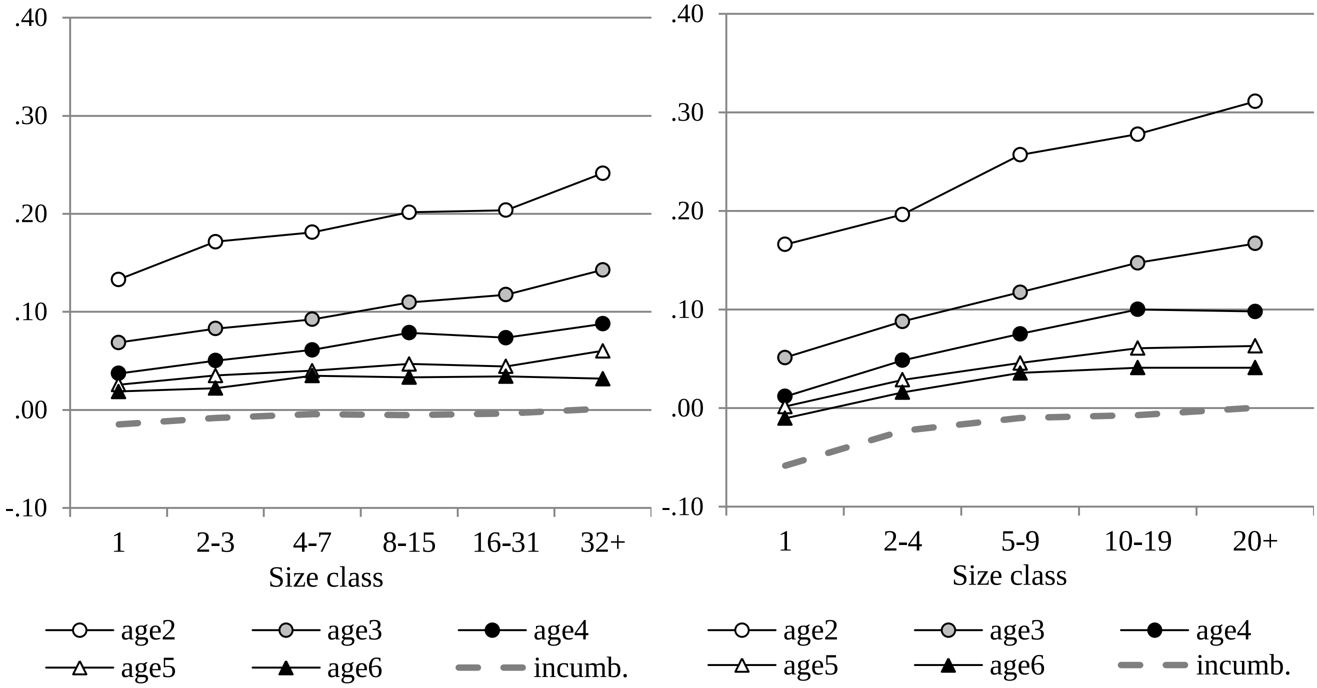


Note: Regression coefficients estimated over the full sample period (2004-2012) using employment weights. Both panels report the same coefficients. Incumbents are firms older than 6 years.

Figure 5: Robustness checks on growth rates of surviving de novo entrants

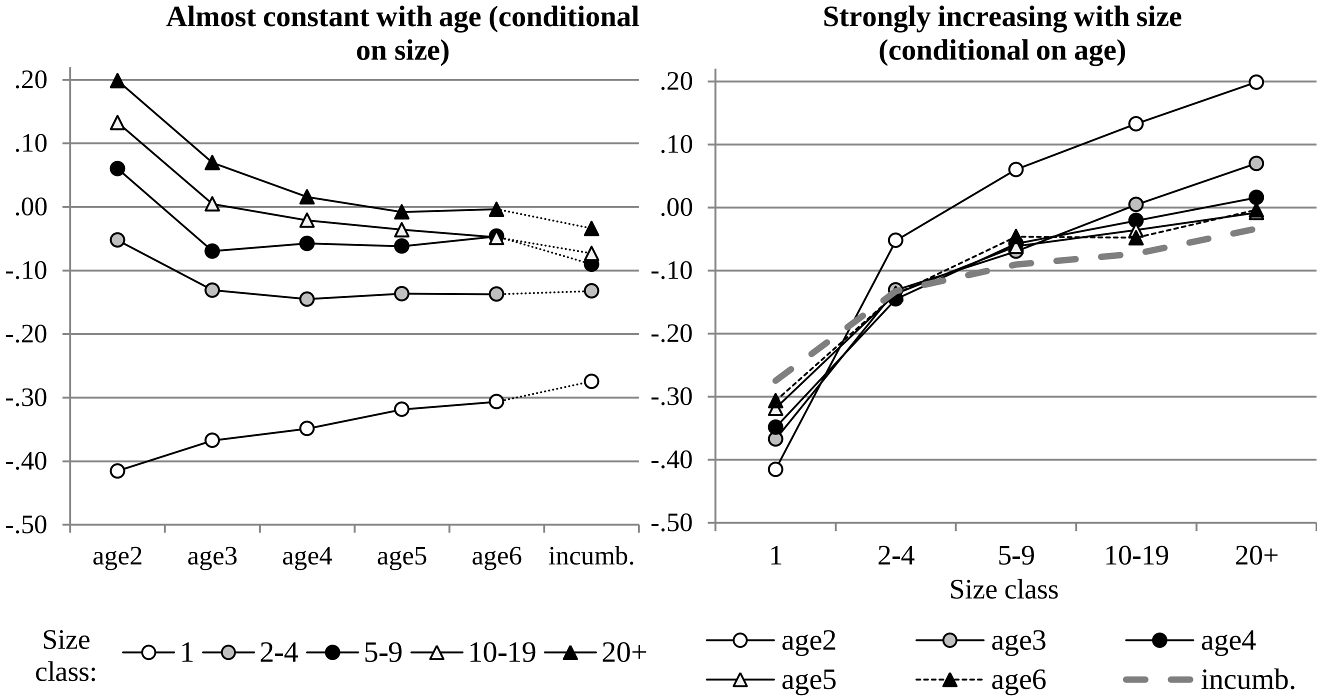
(a) Dynamic size classification

(b) Average of estimates using firm size at *t* -1 and *t*



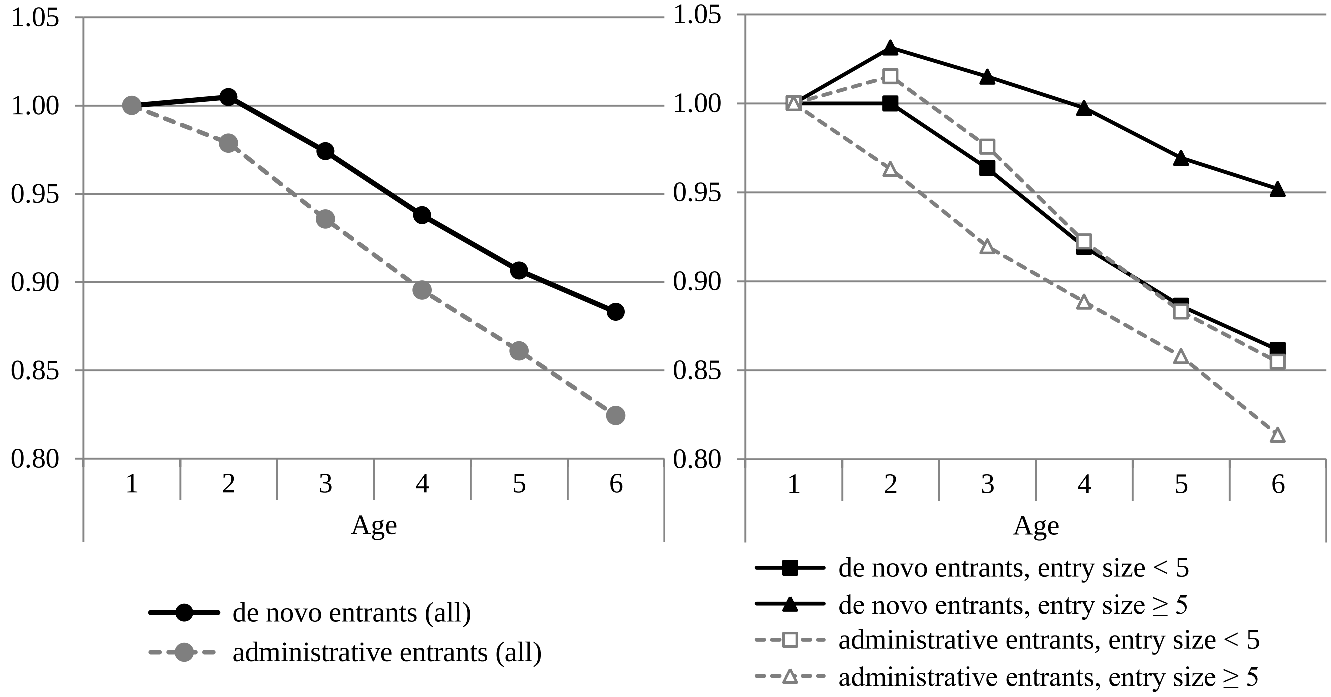
Note: Regression coefficients estimated over the full sample period (2004-2012) using employment weights. Incumbents are firms older than 6 years.

Figure 6: Growth rates of all de novo entrants



Note: Regression coefficients estimated over the full sample period (2004-2012) and all firms (survivors and exiters) using employment weights. Both panels report the same coefficients. Incumbents are firms older than 6 years.

Figure 7: Predicted employment evolution of an entry cohort



Note: Based on regression coefficients estimated over the full sample period (2004-2012). Total employment for each group is normalized to 1 in entry year.

Appendix

Table A.1: Six main industries and NACE Rev. 2 classes

Nace Rev. 2 classes	De novo entrants	
	Firms	Jobs
1. Manufacturing and energy	777	1,996
Section B, C, D, E		
2. Construction	2,730	5,150
Section F		
3. Wholesale and retail trade	4,236	7,497
Section G		
4. Accomodation and food services	2,793	6,570
Section I		
5. Business support services	2,945	5,143
- Freight transport, handling and storage: Nace 49.2, 49.4, 49.5, 50.2, 50.4, 51.2, 52.1, 52.241,52.249;		
- IT programming and services: Nace 62, 63;		
- Central banks, holdings, financial leasing, hedgefunds and auxiliary financial services: Nace 64.110, 64.2,64.3, 64.910, 64.991, 64.992, 64.999, 66;		
- Accounting: Nace 69.2;		
- Head offices: Nace 70;		
- Architecture and engineering: Nace 71;		
- Advertising: Nace 73;		
- Professional and technical support services: Nace 74;		
- Professional rental and leasing: Nace 77.1, 77.3, 77.4;		
- Security: Nace 80;		
- Services to buildings except Cleaning: Nace 81 excl. 81.210 & 81.220;		
- Administrative services: Nace 82;		
- Repair of ICT: Nace 95.1		
6. Mixed business & household services	2,011	3,459
- Passenger transport and transport services: Nace 49.1, 49.3, 50.1, 50.3, 51.1, 52.210, 52.220, 52.230, 52.290;		
- Postal and courier activities: Nace 53;		
- Publishing, Movies, radio and television: Nace 58, 59, 60;		
- Telecommunication: Nace 61;		
- Banks, credit, insurances institutions: Nace 64.190, 64.921, 64.922, 64.92, 65;		
- Real estate: Section L;		
- Legal activities: Nace 69.1;		
- Scientific research: Nace 72;		
- Veterinary : Nace 75;		
- Rental and leasing of household goods: Nace 77.2;		
- Travel agencies: Nace 79;		
- Repair of household goods: Nace 95.2;		
- Personal service activities: Nace 99		
Total	15,492	29,815

Note : Annual averages over the 2004-2012 period of *de novo* entrants and employment in entry year. Firms not in the listed categories are excluded from the analysis, primarily quasi-public sector services and subsidized household help.

Table A.2: Employee flow links by decision rule

	All links	*De alio* entrants	Transfers
1. ID-change	0.57	0.78	0.71
2. Takeover 75%	0.22	-	0.15
3. Takeover 50%	0.01	-	0.02
4. Split-off 75%	0.12	0.18	0.05
5. Split-off 50%	0.01	0.01	0.01
6. Merger of exiters	0.01	-	0.01
7. Break-up into entrants	0.01	0.01	0.01
8. Merger other	0.01	0.01	0.02
9. Break-up other	0.00	0.00	0.01
10. Cluster >= 30	0.03	0.00	0.01

Note: Total sums to one in each column. Annual averages over the sample period.

A link between two different firm identifiers is established if a cluster of at least 5 employees moves from the predecessor identifier in quarter q-1 to the successor identifier in quarter q and one of the following rules is met:

1. ID-change (1 to 1 link)
   cluster is >= 50% of predecessor employment and >= 50% of successor employment
2. Takeover (1 to 1 link)
   Predecessor exits. Cluster is >= 75% of predecessor employment and moves to active successor (no entrant).
3. Takeover bis (1 to 1 link)
   Predecessor exits. Cluster of at least 10 employees is >= 50% of predecessor employment and moves to active successor (no entrant).
4. Split-off (1 to 1 link)
   Successor enters. Cluster is >= 75% of successor employment and moves from surviving predecessor (no exiter).
5. Split-off bis (1 to 1 link)
   Successor enters. Cluster of at least 10 employees is >= 50% of successor employment and moves from surviving predecessor (no exiter).
6. Merger of exiters (n to 1 link)
   We consider the n predecessors as one company, then we apply decision rule 1. All predecessors exit. The sum of the clusters is >= 50% of the sum of predessors' employment and >= 50% of successor employment.
7. Break-up into entrants (1 to n link)
   We consider the n successors as one company, then we apply decision rule 1. All successors enter. The sum of the clusters is >= 50% of predessors employment and >= 50% of the sum of successors' employment.
8. Entrant is created as e merger of parts of firms (n to 1 link)
   Entrant is composed for at least 50% of parts of n predecessors; each part makes up at least 25% of entrant. Successor enters. The sum of the clusters is >= 50% of successor employment, and each individual cluster is >= 25% of successor employment.
9. Exiter is broken-up into parts (1 to n link)
   Predecessor exits and is broken up in at least 2 parts of at least 25%, which continue in separate successors. The clusters represent 50% of the work force of each successor. Predecessor exits. The sum of the clusters is >= 50% of predecessor employment, and each individual cluster is >= 25% of predecessor employment and >= 50% of successor employment.
10. Cluster >= 30 employees (1 to 1 link)

Table A.3: Coefficient estimates underlying the different figures in the text

(a) Figure 3: Exit rates of *de novo* entrants

	Firm size class (employment)				
	1	2-4	5-9	10-19	20+
De novo entrants					
age 2	0.248 (0.002)	0.137 (0.002)	0.110 (0.003)	0.113 (0.004)	0.092 (0.004)
age 3	0.203 (0.003)	0.109 (0.002)	0.096 (0.003)	0.080 (0.004)	0.056 (0.003)
age 4	0.183 (0.003)	0.094 (0.002)	0.070 (0.003)	0.062 (0.004)	0.045 (0.003)
age 5	0.163 (0.004)	0.081 (0.002)	0.056 (0.003)	0.048 (0.003)	0.036 (0.003)
age 6	0.155 (0.004)	0.074 (0.003)	0.042 (0.003)	0.042 (0.004)	0.025 (0.003)
Incumbents					
(age > 6)	0.126 (0.002)	0.056 (0.001)	0.040 (0.001)	0.032 (0.001)	0.017 (0.000)

(b) Figure 4: Growth rates of surviving *de novo* entrants

	Firm size class (employment)				
	1	2-4	5-9	10-19	20+
De novo entrants					
age 2	0.085 (0.003)	0.222 (0.003)	0.279 (0.004)	0.356 (0.006)	0.380 (0.006)
age 3	0.040 (0.004)	0.088 (0.003)	0.122 (0.004)	0.165 (0.005)	0.182 (0.005)
age 4	0.017 (0.005)	0.044 (0.003)	0.082 (0.004)	0.103 (0.005)	0.105 (0.004)
age 5	0.007 (0.006)	0.025 (0.003)	0.050 (0.004)	0.060 (0.005)	0.065 (0.004)
age 6	0.003 (0.007)	0.011 (0.004)	0.038 (0.004)	0.037 (0.005)	0.047 (0.004)
Incumbents					
(age > 6)	-0.023 (0.003)	-0.021 (0.001)	-0.011 (0.001)	-0.009 (0.001)	0.001 (0.000)

(c) Figure 5(a): Growth rates of surviving *de novo* entrants using the dynamic size classification

	Firm size class (employment)					
	1	2-3	4-7	8-15	16-31	32+
De novo entrants						
age 2	0.133 (0.002)	0.172 (0.002)	0.181 (0.002)	0.202 (0.003)	0.204 (0.004)	0.241 (0.005)
age 3	0.069 (0.003)	0.083 (0.003)	0.092 (0.003)	0.110 (0.003)	0.117 (0.004)	0.143 (0.004)
age 4	0.037 (0.004)	0.050 (0.003)	0.061 (0.003)	0.079 (0.003)	0.074 (0.004)	0.088 (0.004)
age 5	0.026 (0.004)	0.035 (0.003)	0.040 (0.003)	0.047 (0.003)	0.044 (0.004)	0.060 (0.004)
age 6	0.019 (0.005)	0.022 (0.004)	0.035 (0.003)	0.033 (0.004)	0.034 (0.004)	0.032 (0.004)
Incumbents						
(age > 6)	-0.015 (0.002)	-0.008 (0.001)	-0.004 (0.001)	-0.005 (0.001)	-0.004 (0.001)	0.001 (0.000)

(d) Figure 5(b): Growth rates of surviving *de novo* entrants; average of results with year *t* and *t-1* size

	Firm size class (employment)				
	1	2-4	5-9	10-19	20+
De novo entrants					
age 2	0.166 (0.004)	0.197 (0.003)	0.257 (0.004)	0.278 (0.006)	0.311 (0.006)
age 3	0.051 (0.004)	0.088 (0.003)	0.118 (0.003)	0.147 (0.004)	0.167 (0.004)
age 4	0.012 (0.005)	0.049 (0.003)	0.075 (0.003)	0.100 (0.004)	0.098 (0.004)
age 5	0.002 (0.005)	0.029 (0.003)	0.046 (0.003)	0.061 (0.004)	0.063 (0.004)
age 6	-0.010 (0.006)	0.016 (0.003)	0.036 (0.003)	0.041 (0.004)	0.041 (0.004)
Incumbents					
(age > 6)	-0.058 (0.004)	-0.023 (0.002)	-0.010 (0.002)	-0.007 (0.002)	0.001 (0.001)

(e) Figure 6: Growth rates of all *de novo* entrants

	Firm size class (employment)									
	1		2-4		5-9		10-19		20+	
De novo entrants										
age 2	-0.416	(0.005)	-0.052	(0.005)	0.060	(0.007)	0.133	(0.009)	0.199	(0.010)
age 3	-0.367	(0.007)	-0.131	(0.005)	-0.070	(0.006)	0.005	(0.008)	0.070	(0.008)
age 4	-0.349	(0.008)	-0.145	(0.005)	-0.057	(0.007)	-0.021	(0.008)	0.016	(0.007)
age 5	-0.318	(0.009)	-0.137	(0.006)	-0.062	(0.007)	-0.036	(0.008)	-0.008	(0.007)
age 6	-0.307	(0.010)	-0.137	(0.006)	-0.046	(0.007)	-0.048	(0.009)	-0.004	(0.008)
Incumbents										
(age > 6)	-0.274	(0.005)	-0.132	(0.002)	-0.090	(0.002)	-0.073	(0.002)	-0.034	(0.001)

Figure A.1: Average size of de novo entrants in the year before event

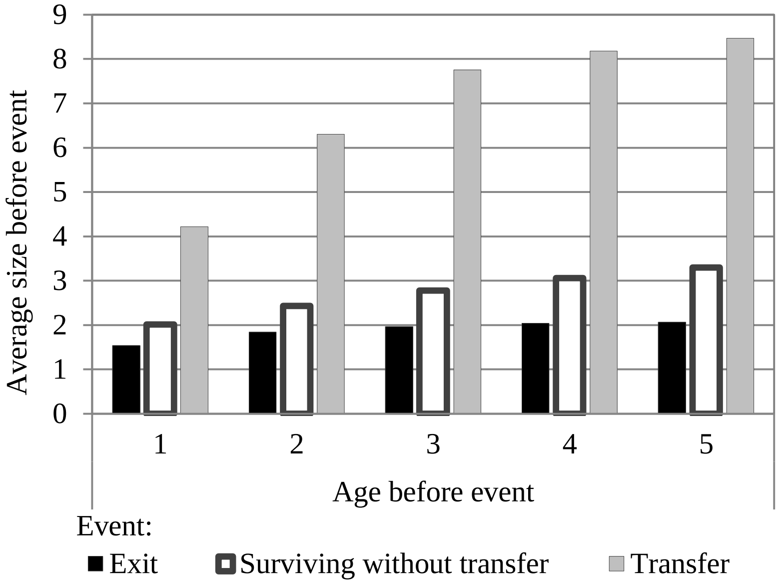
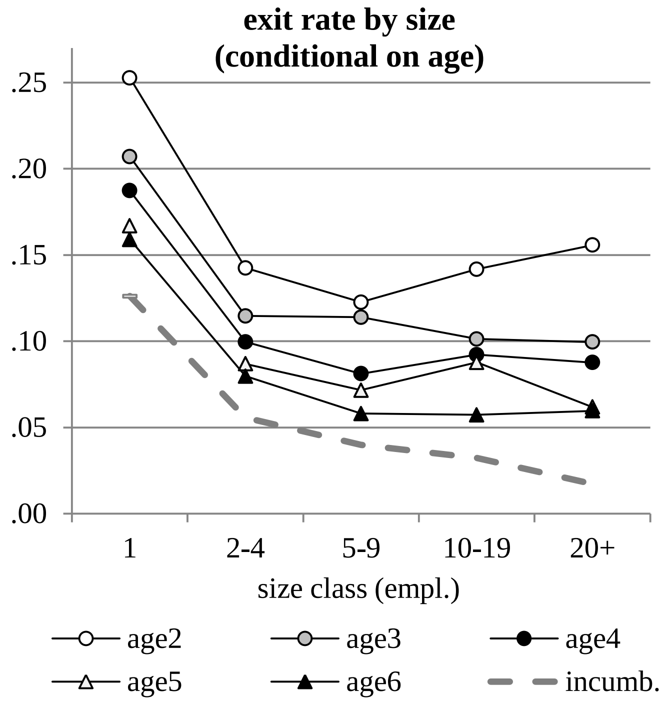


Figure A.2: Exit rates without proper corrections of firm restructuring

(a) *De novo* entrants, not corrected for transfers



(b) *De alio* entrants (also not corrected for transfers)

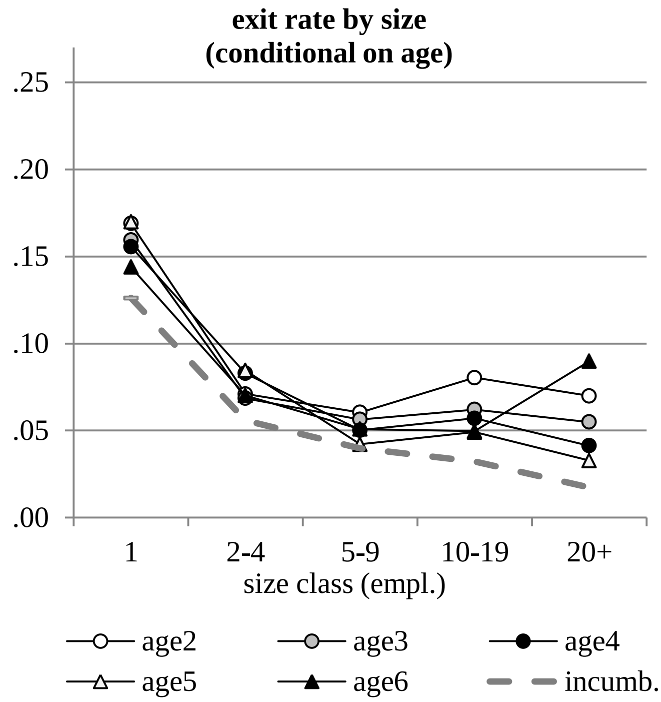
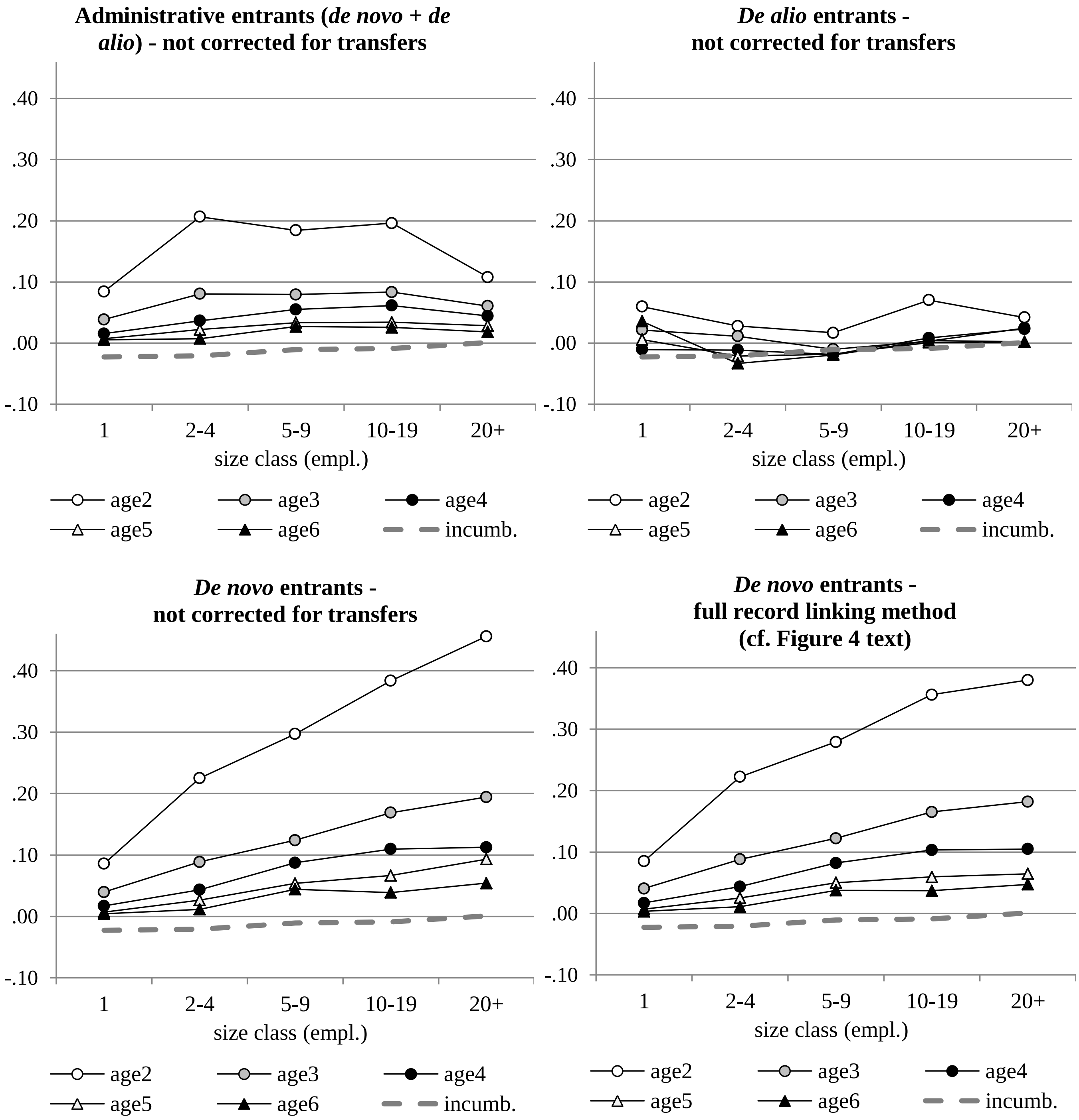


Figure A.3: Growth rates of surviving entrants without proper corrections of firm restructuring



Note: Regression coefficients estimated over the full sample period (2004-2011) using employment weights. Incumbents are firms older than 6 years.

Figure A.4: Growth for surviving de novo entrants, separately for pre and post-crisis entrants

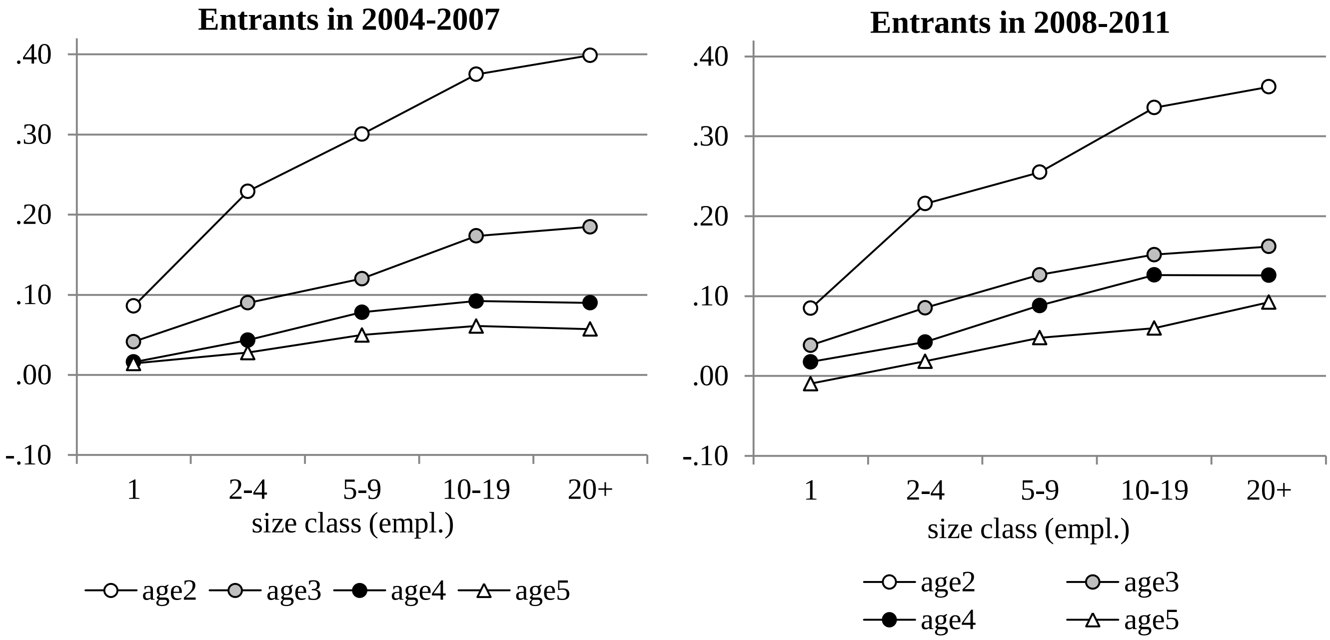


Figure A.5: The growth-size relationship for survivors conditional on age by industry

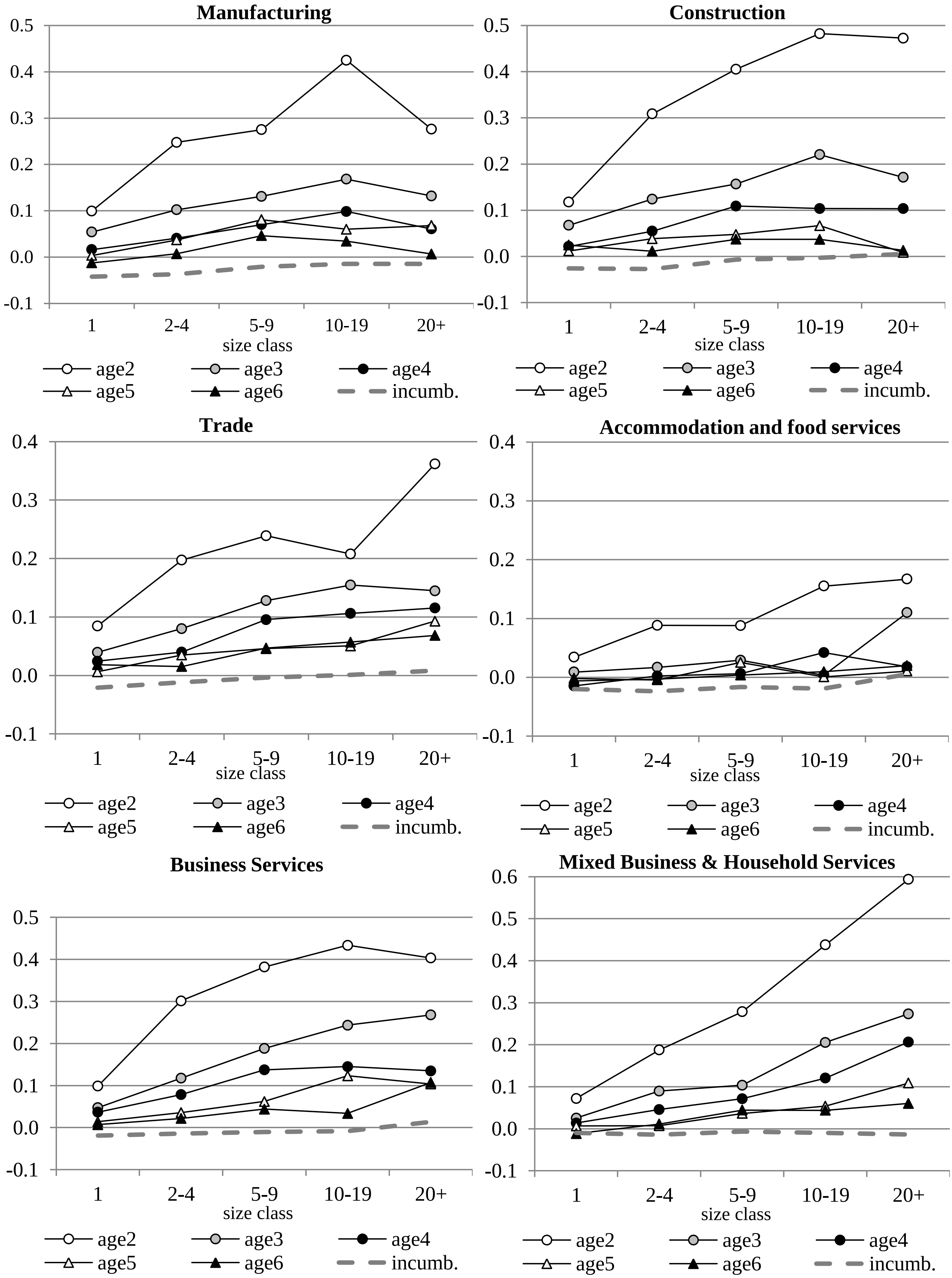


Figure A.6: Evolution of firm size distribution and mean growth rates for de novo entrants