KU LEUVEN
FACULTY OF ECONOMICS AND BUSINESS

Job creation and firm dynamics

Understanding the role of entry, survival and restructuring

Dissertation presented to obtain the degree of Doctor in Economics

by

Karen Geurts

Number 525 2016

Doctoral Committee

Supervisor:

Prof. dr. Johannes Van Biesebroeck, KU Leuven

Committee:

Prof. dr. Eric Bartelsman, VU University Amsterdam

Prof. dr. John Earle, George Mason University

Prof. dr. Joep Konings, KU Leuven

Prof. dr. Otto Toivanen, KU Leuven

Chair:

Prof. dr. Stefan Proost, KU Leuven

Acknowledgements

Writing this thesis would not have been possible without the support and critical feedback of friends and colleagues. I especially want to thank my supervisor Prof. Jo Van Biesebroeck for teaching me with tireless patience how to conduct economic research. His quick and intelligent remarks on numerous preliminary versions of the papers have often taken me by surprise. While I had a job during the week and sent him drafts on Sunday evening, I usually received his detailed comments the same night. From meetings when he offered insights that urged me to start all over again, I left with new energy and inspiration. My sincere thanks to the members of my doctoral committee, Prof. Eric Bartelsman, Prof. John Earle, Prof. Joep Konings and Prof. Otto Toivanen, for their rigorous and generous readings of the papers. The inspiring afternoon when Jo brought these four esteemed scholars together around a table, and they commented extensively on the drafts, was a unique experience in my career as a researcher. The valuable suggestions for improvement they provided have shaped the final chapters presented in this manuscript.

These four years of intense study have been a time of surprising discoveries and unexpected hurdles. I express my gratitude to my mother and father, and all family, friends and colleagues for their encouragement. Some of them have been close witnesses of the long hours of work. I would like to thank them particularly. Michel Lanssens has been a gentle and invaluable support to me throughout this process. He provided me with endless hours of lively conversation on economic history and real life business management. He has been my sounding board, my first reader, and most importantly, a source of joy and friendship. At the faculty, I have been extremely lucky to share the office with Alex Schmitt. As a true *compagnion de route*, he inspired me with many fruitful discussions and sharp insights into economic problems. There have been days this project was particularly demanding. I am very grateful for the friendship and support of An van Dienderen, Gilles Marquenie, Thalia Kruger and Eric Somers. Their cheerful company provided a kind reminder that there is life beyond a PhD.

My sincere thanks also to the staff of the faculty, Karla Vander Weyden, Pascale Kapers, Heidi De Vadder and Annie Vercruysse for their professional and friendly assistance.

Lastly, I want to thank my son, Brand. All these years he waited patiently when I was late again with dinner. When we finally sat together, he just gently asked me whether I did have time to notice the beautiful weather.

Contents

List of Tables

In Appendix

List of Figures

In Appendix

Abstract

In market economies, firms are the main drivers of job creation and job destruction. This thesis focuses on two crucial stages in the life-cycle of firms that involve a large amount of job reallocation: the early years after startup and the period when two firms merge into a single company. The first stage is believed to be a major source of job creation. The second is associated with massive job loss. I investigate whether these common perceptions are supported by empirical evidence, and show that the employment outcomes differ strongly depending on the firm's individual characteristics. To analyze these questions we need data that accurately reflect the dynamics of firms and employment in the economy. This issue is addressed first.

Chapter 1

General introduction

In market economies, firms are the major source of job creation and job destruction. New firms start up, expand, restructure or contract, and eventually exit the market. At each stage they create new employment positions for workers or destroy existing jobs. The magnitude of this job reallocation is enormous. In western economies, about 10 percent of all existing jobs disappear every year and about the same amount of new jobs are created.

The ongoing process of job creation and job destruction is partly driven by macroeconomic changes affecting many firms in similar ways. Technological progress has caused agricultural employment to disappear and has recently driven the rise of service jobs. Aggregate shocks, such as changes in energy prices or a global crisis, lead to recessions with overall job losses across firms, followed by periods of economic recovery in which many firms create new employment opportunities.

But macroeconomic evolutions are only half of the story. A growing body of microeconomic research has documented that aggregate changes explain only part, and in fact very little of the ongoing process of job creation and job destruction. Even within the same sector and in the same period, expanding firms and new firms emerging coexist with contracting firms and firms that fail. During the internet rise in the late 90s, many dot-com companies started and some like Google or Amazon expanded tremendously, but millions of others stayed small or failed. In the once flourishing Belgian textile industry, most firms have declined and eventually disappeared, yet others like Picanol or Van de Velde have restructured their production and are today successful firms providing jobs for many workers around the world.

Individual firm characteristics

The magnitude of dynamism within sectors brings to the fore that firm-specific features are a key element to understand why some firms survive and create jobs, and others destroy jobs or fail. Firms differ in managerial ability, the competences of their workers, the way they invest in new technologies or interact with foreign markets. Each of these elements determine how 'efficient' firms are and how well they are able to successfully compete in the market.

Microeconomic models have developed the frameworks to understand how these differences in efficiency drive a continuous selection process among firms (Jovanovic 1982; Hopenhayn 1992). Even if firms face unforeseeable events they have no control over, efficient firms are more likely to survive and expand than their less efficient counterparts because they are able to produce the same product or provide the same service at a lower cost. This leads to ongoing reallocation of jobs, workers and capital from less to more efficient firms. At the level of the total economy, this selection process is an important source of average productivity increases and welfare growth.

A growing body of empirical studies tries to understand the drivers and implications of this dynamic process by analyzing patterns observed in actual firm-level data. Particular areas of this research investigate dimensions such as the entry and exit of firms, the importance of innovation, the role of market distortions and institutions, and the effects on aggregate output, productivity and employment. In thesis, I focus on two crucial stages in the life-cycle of firms that involve a large amount of job reallocation: the entry of new firms (chapter 3), and the merger of two firms into a single company (chapter 4). The first stage is believed to be a major source of job creation; the second is associated with massive job destruction. I investigate to what extent these common perceptions hold, and how different firm characteristics at these stages lead to different employment outcomes. A prerequisite for analyzing these questions is the availability of good data with reliable information on firm and employment dynamics. Chapter 2 addresses this issue.

This first chapter provides a general introduction to the three topics. It explains why it is important to address these questions and outlines the research background to the subjects. It gives a brief overview of the insights reached so far and of the contributions of my research to previous literature. The specific theoretical and empirical frameworks are discussed in the introductory sections of each chapter.

The three chapters of this thesis are written as stand-alone papers. Since each of them is self-contained and may be read independently, there is inevitable some repetition across the chapters regarding the data and methodological approach.

Main concepts

Before we start, let us briefly clarify the main concepts that we use throughout this thesis.

Firm dynamics is a general term that embraces three processes (Caves 1998): the births and deaths of firms (entry and exit), changes in the size of continuing firms (expansion and contraction), and shifts between enterprises in the control of business units (restructurings).

A job denotes an employment position filled by an employee within a firm (Davis, Haltiwanger and Schuh 1996a). A job can be filled by the same person for a certain period, but when he or she leaves the firm and is replaced by another employee, it is considered as the same job.

One concept remains to be explained. What is a firm? This is the question we turn to next.

Finding the firm
Chapter 2: Empirical measurement of firm dynamics

Firms in economic analysis are generally thought of as profit maximizing organizations that use input factors, such as labor and capital, to produce a certain amount of output, which can be goods or services. Yet unlike individuals, which have a distinct shape and an unambiguous moment of birth and death, the boundaries of a firm, both in space and in time, are less clearly defined.

In a short paper that revolutionized our understanding of the firm, Ronald Coase (1937) defined it as an organization in which market transactions, typically regulated by prices, are eliminated and substituted by the coordination of the entrepreneur. Firms arise when these market transactions can be organized at a lower cost inside a formal organization, i.e. by establishing a long-term contract between economic agents. In the real world, Coase suggested, these contracts are best approached by the legal relationship between an employer and an employee. In this view, the concept of an employer enterprise is a good starting point for the analysis of firm dynamics.

The property rights approach takes a different view of the firm (Grossman and Hart 1986). It defines the firm's boundaries in terms of the ownership of assets (e.g. machines, buildings). When it becomes too costly for two owners, for example a buyer and a supplier, to enforce all the details of the contract, it may be optimal for one party to purchase the other firm and gain control over its assets. The implication of this definition is that the firm can be ultimately traced back to its shareholders.

Williamson (1979) refined the concept of vertical integration of firms by arguing that contractual relationships between economic agents are not that unidimensional. They vary across a set of key characteristics, such as frequency, uncertainty, and required investments. The cost of the transaction along these dimensions will eventually determine whether the contract is governed in a hierarchical relationship within the firm, or by the market. According to this view, firms are best described as governance structures.

Modern firm theory emphasizes that a sharp distinction between intrafirm and interfirm transactions cannot be drawn. Many intermediate forms of organizational structure exists in a continuum between the market and the fully integrated firm, such as franchising, subcontracting, or interfirm networks (Holmström and Roberts 1998). So, where do we start to analyze the firm?

Who creates jobs?

Different research questions require different definitions of the firm. When café Kaminsky changed owner but continued with the same infrastructure, the same employees, and the same relations with its suppliers, should the old and new Kaminsky be considered as a two firms, an exit and a startup, or as one continuing firm? Another example is when Dreambaby was established as an independent legal enterprise but remained under corporate control of Colruyt Group. Was it to be considered as a new entrant in the market of baby articles and as an example of the increasing vertical disintegration of firms, or should it be regarded as the continuing unit of a large corporation diversifying its activities in different product markets?

If our research question is to investigate the impact of firm dynamics on the creation and destruction of jobs, the answer is rather straightforward. It is clearly wrong to define the old Kaminsky as a failing firm that caused job loss for all its workers, and the new Kaminsky as a startup that created new jobs. Since all employees kept their jobs in the bar, we could best approach Kaminsky as a continuing firm. Similarly, the newly established Dreambaby was not a real entrant creating new employment positions, neither was Colruyt a shrinking employer that destroyed all Dreambaby jobs. If any changes in employment occurred at all, it would be best to consider these at the level of Colruyt Group.

The next and more difficult question the researcher faces is, which data can be used to empirically investigate the amount of job creation and destruction that is going on in the real world? While theories of the firm have made major progress in the past 80 years, the datasets researchers can rely on have not. Current empirical research on firm and employment dynamics is mostly based on administrative datasets such as official business registers or social security data. These data are a highly attractive source for analyzing firm-level determinants of entry, exit and growth (Caves 1998) and patterns of job creation and destruction (Davis et al. 1996a). However, the administrative notion of the firm employed in these sources is often at odds with the economic reality. Firms can change administrative ID-code, or legal entities can be separated or merged into one unit, leading to missing links in firm histories over time. Researchers have indicated that this provides us with a distorted picture of firm and employment dynamics (Haltiwanger, Jarmin and Miranda 2013) and that it hampers comparative analysis (Bartelsman, Scarpetta and Schivardi 2005).

Employer-employee relations

The first paper of this thesis addresses the question how to transform administrative datasets into a reliable source for the analysis of firm and job dynamics. Reflecting the real-world firm suggested by Coase, I argue that the employer-employee relationship is a key element to identify the firm and its impact on job creation and destruction. A so-called employee-flow method is developed, pioneered by Baldwin, Dupuy and Penner (1992), to trace each firm's individual history. The method follows one main input factor of the firm, its workforce, to identify the firm's point of entry, exit, and its changing structure over time. I contrast this approach with another method commonly used in Europe and the US, which uses a complex set of characteristics to identify firm histories (Eurostat-OECD 2007; Clayton and Spletzer 2009). I conclude that this traditional method leaves the researcher with an obscure definition of the firm and its boundaries in time and space.

Turning to actual statistics, I show that the employee-flow method is generally more effective for obtaining reliable estimates of job creation and destruction and of firm-level dynamics. The method is preferable on other grounds as well. It is based on algorithms that can be standardized across countries and enhance international comparability of results. It is also a powerful tool for identifying changes in the firm structure which researchers often want to study as events of economic importance, such as mergers and acquisitions, spin-offs, and other changes in the control structure of firms.

Expanding the firm

Chapter 3: Post-entry dynamics of de novo entrants

New and young companies are often considered as the primary source of job creation and growth. As an 'answer to challenges brought by the gravest economic crisis in the last 50 years', the European Commission declares that 'to bring Europe back to growth and create new jobs, we need more entrepreneurs' (European Commission 2013).

The idea that new entrepreneurial activity provides a fundamental impulse to economic growth is commonly attributed to the Austrian-born economist Joseph Schumpeter. In his view, technological innovation, as the driving force of economic growth, is accompanied at the micro level by a process of creative destruction. With this term, he denoted the disruptive process of transformation that 'revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one'. For Schumpeter the 'innovative entrepreneur' is the disruptive force. As opposed to the 'imitator', it is the 'leader', who ensures 'fundamental improvement, is able to break away from routine and destroy existing structures' (Schumpeter 1942).

With the passage of time, Schumpeter's distinctive attribute innovative has faded, and the view that new entrepreneurs are a major driver of employment growth has gained popularity. Empirical analysis for various countries, however, shows mixed evidence. New firms may stimulate efficient reallocation of production factors across firms and thus increase aggregate productivity (Foster, Haltiwanger and Krizan 2001). But the displacement of less productive jobs in other firms may lead to a negative impact on net employment growth, albeit in the short run (Fritsch and Mueller 2004). Moreover, ample evidence shows that most new firms fail short after entry or stay very small, hardly contributing to job creation at all (Geroski 1995). Recently, research attention has therefore shifted towards the heterogeneous characteristics of new firms. One finding of these studies which has received major public attention, is that a small set of rapidly growing young firms contribute disproportionally to job creation (Haltiwanger et al. 2013). Stretching this conclusion, some now argue that policy should not try to encourage as many startups as possible, but to stimulate those high-potential entrants that are most likely to become the next Apple or Amazon. In short, do we need more mice or gazelles?

How many new jobs young firms create will ultimately depend on the characteristics of the individual firms both at startup and post entry. It depends on their absolute importance, reflected in the initial firm size distribution, and on changes in the distribution over time, captured by firm exit and growth patterns. The second paper of this thesis reassesses some facts about these two patterns.

How do firms enter?

Two distinct theoretical views on the firm size distribution at entry prevail. The first tries to explain heterogeneity among firm size at startup as observed in many empirical studies. Lucas (1978) features a dispersion of managerial skill in the population. High-skilled individuals self-select into entrepreneurship and choose their firm size optimally upon entry. Another explanation is given by Evans and Jovanovic (1989) who take into account that entrants may face liquidity constraints. Heterogeneity in entry size reflects that some firms are more financially constrained. In the second view, there is no difference in size at entry. The passive learning model of Jovanovic (1982) assumes that new firms enter with an innate efficiency level which they only discover from operating in the market. Initially, they have the same beliefs about this and all enter at the same size.

Remarkably, and in contrast to most previous studies, our results are very much in line with the passive learning model. Focusing on *de novo* entrants, we find that the size distribution of new firms is confined to very narrow range of small size classes. We show that this result strongly depends on the identification of truly new firms in the data, as explained in the previous chapter. Studies that cover a sample with many large firms already at entry, are most probably including established firms that are misclassified as entrants.

Although this seems a trivial data problem, it is not. The narrow size range we observe at entry has two important implications. First, after removing misclassified entrants from the data, we find that the initial contribution of new firms to job creation is actually very low. In Belgium, they represent a mere 1.5 percent of total employment, in contrast to a multiple of this share reported in many other studies. The second implication concerns the post-entry growth patterns. Do small young firms grow faster than larger ones, supporting the view that policy should focus on the wide set of micro-firms to stimulate job creation? Or do larger young firms have higher growth rates?

How do young firms expand?

Empirical evidence has long supported the first view. Evans (1987a) suggested that growth rates of young firms of the same age are negatively related to size. Theoretical models have tried to explain this empirical observation by rationalizing how heterogeneity among firm size at startup leads to heterogeneous firm growth paths. In Evans and Jovanovic (1989), for example, liquidity constraints force some firms to enter below their optimal size. Relying on retained earnings to expand, the smaller, constrained entrants would then grow faster and to some extent catch up in size with larger entrants.

Recently, however, Haltiwanger et al. (2013) concluded that there is no systematic relationship between firm size and growth, and that the observed relation depends on the size methodology. Using their preferred methodology even would suggest a positive growth-size relationship among firms of the same age. Our results strongly confirm the latter. We find that among young firms of the same entry cohort, larger ones grow faster than smaller ones. Moreover, we show that this pattern is robust to alternative measurement methods.

Both the narrow initial size distribution and the positive size-growth relationship we find are supportive of constraints affecting firms following their entry decision rather than before. The reasoning of the passive learning model (Jovanovic 1982) is that firms gradually discover their own efficiency level by operating in the market. Firms that learn they are more efficient grow and survive, while the inefficient decline and exit. In the model, these size adjustments are made instantaneously. In the real world, however, young firms may face severe credit, hiring, or regulatory constraints (Cabral and Mata 2003). For more efficient firms wanting to expand, this limits growth in the first years. Their current size will be below their desired size and they will need several years to grow into their optimal size. This leads to higher growth rates for larger firms, until adjustment is complete and growth becomes independent of firm size.

Mice and gazelles

How do these firm-level patterns translate into aggregate numbers of job creation? For Belgium, we find that *de novo* entrants of a given year represent about 30 000 new jobs, which is 1.5 percent of total private employment. Five years after entry, only 1 percent of them have expanded beyond 20 employees. About half of all entrants have already failed by that age and most others have remained very small. Does this imply that job creation by young firms is mainly driven by a limited set of rapidly growing companies? As long as the new Apple or Amazon is not among them, it is probably not. Total employment of an entry cohort has dropped below its initial level of 30 000 jobs by age 5 and is still mainly located among the smallest firms. Fast-growing firms that expanded beyond 20 employees represent 20 percent of these jobs, which is a disproportionate, but after all a modest share.

A straightforward conclusion could be that small young firms create more jobs. A more provocative one is that the next Apple or Amazon may have failed. In global markets, innovative entrants often have only a narrow window of opportunity to occupy a market niche. If, as we described, constraints limit their growth in the first years and scaling-up happens too slowly, a firm risks coming too late and be shut out of the market by early movers. Recent evidence has indeed suggested that fast-growing firms experience the greatest constraints to growth (Brown, Earle and Morgulis 2015). If this is the case, policy could play a role in fostering job creation by making sure that adjustments to firm size after entry are easy to make. Understanding the implications of such constraints on early growth is a promising area for future research.

Restructuring the firm
Chapter 4: Employment impact of takeovers

Firms usually expand gradually by hiring one or a few additional employees at a time. Yet at some point, they may increase drastically in size by taking over or merging with another firm. Mergers and takeovers are driven by such motivations as increasing market power, acquiring innovative technology, or creating gains to shareholders. But when Inbev took over Hoegaarden, Concentra and Corelia were merged, or IJsboerke was acquired by Glacio, one question fascinated the public: how many jobs will be lost?

Every year, more than 6 percent of all employees in the Belgian private sector are working in a company that is involved in a takeover.[1] If takeovers do

[1] This number even refers to domestic takeovers only.

significantly reduce the firm's demand for labor, the consequences for aggregate employment may be considerable.

Expected employment outcomes

From a theoretical perspective, the fear of job loss can readily be justified. A purely anti-competitive merger reduces output and thus employment. An important strand of the literature has focused on such oligopoly behavior of firms (Shapiro 1989). By eliminating competition between the two companies, the integrated firm may exploit its market power and substantially increase the price of its product at the same time reducing its output.

A merger or takeover that increases labor productivity without changing the level of output, will reduce employment as well. Theory provides ample reason to assume that the vertical integration of firms leads to gains in labor productivity (Lafontaine and Slade 2007). Productivity increases can for instance be realized by production cost savings, arising from economies of scale or from efficiently reallocating production and workers across the integrated firm.

An additional motivation for employment reductions following takeovers is presented by Shleifer and Summers (1988). The authors argue that if a new management is appointed with less ties to the workforce of the acquired firm, it will be less reluctant to renegotiate existing labor contracts. This 'breach of trust' between the management and the employees may lead to substantial layoffs.

Despite the insights offered by merger theory, empirical research has not found clear evidence that mergers and takeovers, as a general rule, reduce output or enhance productivity. Gugler, Mueller, Yurtoglu and Zulehner (2003), for example, analyzing the effects of mergers and takeovers around the world over the past 15 years, find that 29 percent of the firms increase efficiency post-merger, but that an equal proportion decrease efficiency. They also find that output-reducing mergers account for about half of the population, while the other half increases output. These results suggest that the outcomes for employment will be highly ambiguous. Moreover, certain types of takeovers may increase productivity without saving in labor costs. One example are acquisitions targeted at small innovative firms, creating synergy gains for both parties. The takeover creates an opportunity for the acquirer to incorporate new technologies and highly specialized personnel, while it provides new resources for the target to finance its technological developments.

What complicates the prediction of merger outcomes even more, is that actual takeovers are often driven by other motivations, which do not fit into the model of a profit-maximizing firm. The free cash flow theory, for example, describes why managers who aim at increasing the resources they control, may choose to expand the firm beyond its optimal size, with detrimental impacts on the firm's performance (Jensen 1986).

Observed employment outcomes

Given the lack of strong predictions that can be derived from the literature, does empirical research provide more explicit support for the alleged job loss associated with mergers and takeovers? The evidence so far is rather weak. Quantitative studies have mainly focused on small subsets of takeovers, such as foreign acquisitions or takeovers by listed firms. Even within these specific populations, the employment effects strongly differ.

One set of studies investigate the employment impact on the target firm only, and mainly focus on domestic firms that are acquired by a foreign owner. Overall, the results indicate that foreign acquisitions have no or small negative effects on employment growth of the domestic plant, and that the impact varies greatly across sectors (Girma and Görg 2003; Lehto and Böckerman 2008).

However, looking at employment changes in the target plant is only half the picture. Takeovers may also affect employment in the acquiring firm, and jobs may be relocated across the integrated company after the merger. When Inbev took over Hoegaarden, the entire production was replaced to Jupille. New investments were made in this plant and new jobs were created. Merely taking into account the jobs that were lost in Hoegaarden, would underestimate the employment impact of the takeover.[2]

A more consistent approach, therefore, is to consider employment changes in both the target and the acquirer. Conyon, Girma, Thompson and Wright (2002) and Gugler and Yurtoglu (2004) have adopted this by estimating the employment impact at the level of the combined entity. These studies, which exclusively focus on takeovers by listed companies, find that takeovers in European countries lead to significant workforce reductions, but they observe no adverse effects in the U.S.

[2] Only two years later, Inbev moved production back to Hoegaarden because the customers disliked the taste of 'de Witte' brewed Jupille.

Mixing different colors

A drawback of the combined-entity approach is that it disregards that takeovers are combinations of two firms with different characteristics before the merger. A large pharmaceutical company taking over a small high-growth IT firm, will reflect a different merger motivation and presumably have a different employment impact than when a medium publishing company merges with a medium retail bookseller. Adding a touch of green to blue is unlikely to produce the same result as mixing red and yellow. Yet the combined-entity method treats them as similar events.

In the third paper of this thesis, we refine the combined-entity approach by taking into account that the characteristics of both the target and the acquirer, and the specific combination between the two may affect the decision to engage in a takeover and subsequent employment growth. In particular, we consider such features as pre-merger size, previous growth, industry, and the corporate structure of the two firms. As a counterfactual for the takeovers, we use pairs of firms with the same combined pre-merger characteristics.

We also use a more comprehensive set of takeovers than previous studies. Our sample consists of 2200 domestic takeovers in the Belgian private sector, which are identified as two independent employer firms that merge into a single legal unit. This setting enables us to explicitly concentrate on the employment effects of merging separate workforces into a larger entity.

Our results indicate that takeovers have a small negative impact on employment growth of the merged entity, which is mainly attributed to takeovers undertaken by small acquirers. For large acquirers we find substantial variation in post-merger employment growth suggesting that workforce rationalizations are not the dominant motivation for takeover activity. In particular, we find suggestive evidence that takeovers targeted at high-growth firms have a positive impact on firm employment growth.

Chapter 2

Longitudinal firm-level data: problems and solutions

Abstract

Empirical measures of firm and employment dynamics based on administrative datasets are biased due to missing links in the longitudinal observation of firms. This paper presents a systematic overview of the problems and evaluates two prevailing solutions. We quantify the biases in a set of widely used empirical measures and show which estimates are most sensitive to missing linkages. The biases are found to be especially large in the size distribution of entrants and exits, in firm-level growth estimates for medium and large firms, and in job reallocation measures. We show that an employee-flow linkage method is more effective in reducing bias than a traditional link method often used by statistical agencies. A consistent approach is developed for imputing firm-level growth measures of linked firms. The analysis is carried out using a longitudinal dataset for Belgium and discussed from an international perspective.

JEL Codes: C81, J23, L11
Keywords: Firm dynamics; Job creation and job destruction; Firm microdata; Linked employer-employee data; Firm linkage

---

This chapter is published in *Small Business Economics*, 2016, 46:3, p. 425-445. I am grateful for the valuable advice of Jo Van Biesebroeck. I am also grateful to the Belgian National Social Security Office (NSSO), in particular Peter Vets for collaboration on the development of the employee-flow method, and to Statistics Belgium, in particular Youri Baeyens and Antonio Fiordaliso for providing firm linkages created for the Eurostat Structural Business Statistics. I also thank Malcolm Brynin and Alex Schmitt for helpful suggestions, and three anonymous referees for valuable comments. Financial support from ERC grant No. 241127 and KU Leuven Program Financing is gratefully acknowledged.

2.1 Introduction

Large-scale administrative datasets in which individual firms or establishments are observed over a long period of time are increasingly used in empirical research on firm and employment dynamics. The data are an attractive source for investigating firm-level determinants of entry, exit and growth (Caves 1998; Dunne et al. 1988, 1989; Wagner 2007) and patterns of job creation and destruction (Davis et al. 1996a). A well-known but major problem with these data is that missing links in individual firm histories lead to bias in the measurement of firm turnover and job reallocation (Haltiwanger et al. 2013). When, for example, the administrative ID number of a firm is changed, the firm is observed as an exit and a new entrant instead of as a continuing firm. Changes in the firm's structure, such as mergers or split-ups, create additional difficulties in identifying firm histories over time. In response to these concerns, statistical agencies have invested in the development of record linking methods to improve the data for research. In the absence of comparison, the reliability of the revised datasets is generally taken for granted by users. Moreover, it is often unclear to researchers which empirical measures are especially sensitive to linkages problems and require good longitudinal data to be consistently estimated.

This paper aims at providing guidance to researchers that use administrative datasets to investigate firm and employment dynamics. It contributes to previous studies by presenting a systematic overview of both the linkage problems and the solutions. First, it evaluates the size and the direction of the biases created by missing firm linkages in a series of widely-used empirical measures. We consider entry and exit indicators, firm-level growth estimates, and the mean and annual variance of job creation and destruction rates. Second, it compares the performance of two prevailing linkage methods in reducing these biases. A novel approach is adopted to address these questions by using reference measurements based on a benchmark dataset. The empirical analysis is carried out using a longitudinal dataset of Belgian firms. The methods and findings are discussed from an international perspective. While this study focuses on the firm level, it provides insights that can also be useful for establishment-level analysis, which is sensitive to similar measurement problems.

Statistical agencies in Europe and the U.S. often apply traditional record linking methods to address longitudinal linkage problems in administrative data (Eurostat-OECD 2007; Clayton and Spletzer 2009). Missing firm linkages are identified with probabilistic matching techniques and supplementary data sources such as surveys or other administrative registers. An alternative approach, which in this paper is called the employee-flow method, was pioneered by Baldwin et al. (1992) for Canada and has been adopted in some other countries. The method relies on linked employer-employee data and traces one key input factor of the firm, the stock of individual employees, to track changes in firm ID numbers and in firm structure. This paper discusses the general strengths and weaknesses of both linkage methods, and applies them to a longitudinal dataset of Belgian employer firms in order to evaluate how well they perform in improving empirical estimates. The methods for Belgium have been developed in line with current international practice and can be considered as illustrative examples of the two linkage approaches. A third linkage method, adopted by the U.S. Census Bureau, is not evaluated in this paper. It takes the traditional method one step further and uses also firm-establishment links to improve longitudinal linkages (Jarmin and Miranda 2002).

The impact analysis of missing linkages on empirical measures of firm and employment dynamics is carried out as follows. Starting from an administrative register of Belgian employer firms from 2003-2012, two improved versions of the dataset are obtained by applying the traditional and employee-flow linkage methods separately, and a third version is constructed by using all linkage information provided by the two methods combined. The latter serves as a benchmark for evaluation. Comparing empirical estimates based on different versions of the dataset, we first investigate which measures are especially sensitive to missing linkages and quantify the size of the biases. Second, it is evaluated how well each of the two linkage methods perform in reducing these biases.

The analysis reveals that missing linkages are strongly increasing in firm size. The implication for aggregate measures is that they lead to large overestimations of job flows and a relatively small bias in firm turnover. This result is consistent with fragmented evidence for other countries. The impact is most obvious for entry and exit measures, where the analysis reveals that most medium and large entrants and exits are actually continuing firms that are misclassified after an ID change or firm restructuring. Using improved firm linkages almost completely shifts the mass of employment at entry and exit towards the smallest firms and reduces total job creation by entry and job destruction by exit by about half. Missing links in individual firm histories further have disturbing consequences for firm-level estimates by age and size, since they lead to misclassifications of older as younger firms and of successful firms as exits. As an example, we show that firm-level growth estimates based on unedited data are strongly underestimated, especially for larger firms.

An optimal firm-level research dataset with reliable longitudinal information is obviously obtained by exploiting information from different linkage methods. If only one method is used, we find that the employee-flow method is generally more effective. Empirical measures based on data improved by this method are close to the benchmark results, while the traditional method reduces most biases by only half. The traditional method is however more appropriate for the study of the smallest firm size classes, where employee-flow linkages are absent by construction. This paper further argues that the employee-flow method is preferable on other grounds as well, such as its use of an economically meaningful definition of firm continuity, its potential for international comparability, and its wider application in the field of firm dynamics.

Section 2.2 discusses the longitudinal linkage problems and the solutions offered so far. Section 2.3 describes the data used in this paper and provides the technical background on the two linkage methods that are evaluated. It also explains how linked firms are reclassified, and proposes a consistent approach for imputing employment growth of these firms. Section 2.4 provides background statistics on linked firms to facilitate interpretation of the results. Section 2.5 presents the results of the impact analysis of linkage errors on empirical measures. Section 2.6 concludes.

2.2 Longitudinal linkage problems and solutions

It is well-recognized that firm-level administrative datasets suffer from longitudinal linkage problems which have disturbing consequences for the empirical analysis of firm and employment dynamics (Baldwin et al. 1992; Vilhuber 2008; Haltiwanger et al. 2013). The problems stem from the fact that the data reflect firm entry and exit, and firm expansion and contraction, from an administrative viewpoint which often does not correspond to the economic reality one wants to investigate. A first linkage problem is created by changes in the administrative firm identification number. A new ID number may be assigned when the ownership or legal form changes, or firms may re-register as a new company after an internal restructuring, for tax optimization or liability avoidance. The firm will then be observed twice, once as an entrant and once as an exit, even if it simply continues its activities with a new ID number. Such 'spurious' entrants and exits introduce an upward bias in measures of firm and employment turnover (Davis et al. 1996a). Changes in the firm's structure brought about by mergers, takeovers or split-offs create additional longitudinal linkage problems. They lead to creations and closures of firm identification numbers that are clearly different from *de novo* firm entry and exit by failure (Dunne et al. 1988; Baldwin and Gorecki 1987). They also lead to administrative transfers of employees between firm ID numbers which appear as shocks to firm-level employment in the raw data. This further inflates job reallocation measures (Pinkston and Spletzer 2002). Overestimation of aggregate measures of firm turnover and job reallocation is one empirical problem. Several authors have pointed out that missing longitudinal linkages also cause distortions in firm-level measurements. They introduce a size bias in firm-level estimates of entry, exit and growth (Geurts and Van Biesebroeck 2014), lead to misclassifications of the firm age (Haltiwanger et al. 2013), and hamper comparative analysis of firm demographics (Bartelsman et al. 2005; Vilhuber 2008).

To improve the data for statistical and research purposes, statistical agencies have invested in the development of longitudinal business databases. Traditional record linking techniques are often used to identify missing links between firm identification numbers (Abowd et al. 1999; Bycroft 2003; Eurostat-OECD 2007; Clayton and Spletzer 2009).[1] These methods primarily rely on supplementary data sources such as surveys and other administrative registers with information on firm demography, ownership changes or M&A activity. Although such sources provide valuable additional information, firm changes that are not registered remain out of scope. Therefore, the link procedures are usually complemented by probabilistic matching techniques which exploit similarities in partial firm identifiers, such as name, address, or industry code to establish links between ID numbers of the same firm.

At the same time, several countries have developed an employee-flow method for the improvement of longitudinal firm linkages (Baldwin et al. 1992; Benedetto et al. 2007).[2] This method takes an entirely different approach to identify missing

[1] Abowd et al. (1999) present a linkage method for France; Bycroft (2003) for New Zealand, and Clayton and Spletzer (2009) describe the longitudinal linkage method applied by the U.S. Bureau of Labor Statistics. In Europe, linkage methods adopted by national statistical agencies have led to general Eurostat-OECD recommendations on firm record linking (Eurostat-OECD 2007).

[2] One of the first institutes to implement an employee-flow method has been Statistics Canada (Baldwin et al. 1992), where it is still used for the construction of the National Accounts Longitudinal Microdata File (Dixon and Rollin 2012). Employee-flow methods are

linkages. While traditional methods retrieve information on firm continuity from a complex set of partial firm characteristics, the employee-flow method uses one key input factor of the firm, the workforce, to trace the firm's individual history and changing structure over time. The reasoning is as follows. When a firm changes ID number but continues its operations, one of the main production factors, the stock of individual employees, is likely to remain largely the same. Continuity of the workforce from one period to the next can thus be used to detect changes in firm ID numbers. Similarly, mergers, takeovers or split-offs of firms will be reflected in a merge or division of workforces. Employee-flow methods make use of linked employer-employee data to implement this workforce-based characteristic of firm continuity. Large clusters of employees that appear to 'move' from one firm ID number to another are used to signal changes in ID numbers or in the firm structure.

A third linkage approach is adopted by the U.S. Census Bureau for the creation of a longitudinal establishment and enterprise database (Acs and Armington 1998; Jarmin and Miranda 2002). A similar method does not exist for Belgium and is therefore not evaluated in this paper. In addition to using surveys and probabilistic matching, as in the traditional approach, the U.S. Census Bureau also exploits information provided by the link between the firm and the establishment ID. One of the advantages of this method is that the longitudinal histories of establishments can be used to identify true entry and exit of the controlling firm as well as changes in the firm's structure. Establishments that change ID number are in turn linked by a probabilistic matching procedure. Haltiwanger et al. (2013) highlight that this approach allows for measuring job creation and destruction at the firm level that abstracts from establishments whose parent firm changes. In the same spirit, Mata et al. (1995) and Baldwin and Gorecki (1987) use the longitudinal identifier of the parent firm to distinguish between *de novo* plant entry and new plants created by established firms.

Although this paper focuses on the firm level, it provides useful insights for establishment analysis as well. Establishment data equally suffer from missing links in the longitudinal registration of units due to ID changes, split-ups or mergers of establishments, which give rise to similar biases in turnover and job

---

also used for the construction of longitudinal employer databases in Denmark (Albaeck and Sorensen 1998), Finland (Korkeamäki and Kyyrä 2000), Sweden (Persson 2004), Italy (Contini et al. 2007), Belgium (Geurts et al. 2009); and Germany (Hethey and Schmieder 2013).

flow measures at the establishment level. The problems are often addressed with similar linkage methods as the ones discussed in this paper.[3]

From a methodological viewpoint, both the traditional and employee-flow linkage methods have strengths and weaknesses. The main disadvantage of the employee-flow method is that it is unsuited to capture links among the smallest firms, while the traditional method covers all size classes. The employee-flow method, however, has several features that make it an attractive method for research on job flows and firm dynamics. First, it directly implements an economically meaningful definition of firm continuity, and by extension of firm entry and exit. Continuity of one of the firm's key production factors, the stock of employees, is used to identify firms that operate continuously but change identification number or firm structure. The traditional method deduces information on longitudinal firm histories from a complex set of partial firm characteristics. It partly depends on the specific notions of firm continuity that are used in supplementary surveys or administrative registers. Moreover, ID changes induced by firms themselves for tax evasion or other reasons are unlikely to be reported in other sources. Probabilistic matching helps to identify additional linkages, but major changes in discriminating identifiers, e.g. name or telephone number, strongly reduce the probability of a positive match. Second, the employee-flow method effectively captures changes in the firm's structure, such as mergers, take-overs or split-offs. Researchers often do not simply want to neutralize these events, as we do in this paper, but to study them as events of economic importance.[4] A third advantage lies in international comparability. Traditional link methods make use of supplementary data sources which differ widely between countries. The employee-flow method, by contrast, is based on linkage algorithms that follow general rules and can be standardized across countries.

From an empirical viewpoint, the strengths and weaknesses of the two methods are less obvious. Both traditional and employee-flow methods are found to remove a substantial amount of 'spurious' firm and employment turnover from the data (see for example Pinkston and Spletzer 2002; Benedetto et al. 2007). Empirical dynamics measures based on improved longitudinal data thus more

[3] See for example Abowd et al. (1999) who describe a traditional link method for French establishment data, and Hethey and Schmieder (2013) who use an employee-flow method for German establishment data.

[4] Linked employer-employee data have further been used to analyze a wide variety of labor market issues. For a comprehensive overview see Abowd and Kramarz (1999).

accurately reflect the true dynamics in the economy. However, to our knowledge, no serious attempt has been undertaken so far to evaluate bias in empirical measures after implementation of the record linking methods, or to determine which linkage method is more successful in avoiding bias in the measures.

2.3 Data and methods

The register of Belgian employer firms that is used in this paper is particularly well suited to investigating these questions. First, the two record linking methods that we apply to the dataset have taken advantage of the development of similar methods in other countries and can be considered as illustrative examples of the traditional and the employee-flow approach. The traditional method was developed by Statistics Belgium in line with the OECD-Eurostat recommendations on firm record linking (Eurostat-OECD 2007). These guidelines aim at the construction of harmonized business registers and statistical indicators on firm demography. The employee-flow method was developed by the National Social Security Office in collaboration with the University of Leuven (Geurts et al. 2009) and builds on similar examples in other countries, in particular Canada (Baldwin et al. 1992), Sweden (Persson 2004), and the U.S. (Benedetto et al. 2007).[5] Further, the initial longitudinal firm linkages present in the Belgian firm register are relatively consistent: firm identification numbers generally do not change after a change in ownership or legal form, while this is often mentioned as one of the main reasons for longitudinal linkage errors in other countries. The benchmark results that are obtained after applying the two linkage methods combined can therefore be considered fairly accurate estimates of firm and employment dynamics in the economy.

The level of analysis used throughout this paper is the firm. The firm (or 'enterprise') is the basic statistical unit for business demography statistics in the European Union, and data collection is harmonized across countries (Eurostat-OECD 2007). The firm corresponds to the smallest enterprise unit which benefits

[5] The methods has been developed for the construction of the DynaM longitudinal employer database. The database is designed to track changes in employment both at the macro and the firm level and to produce annual series of gross job gains and losses statistics in Belgium. See www.dynam-belgium.org

from a certain degree of autonomy in decision-making.[6] In Belgium, as in most European countries, this corresponds to the enterprise entity at the national level, which may include different establishments. 90 percent of the firms in our dataset are single-establishment firms. Some studies, especially for large countries such as the U.S., use the establishment as unit of analysis. In smaller countries, this level of observation is less suitable for job flow analysis, since establishments of the same firm are located at short distance from each other and people easily commute to different work locations. Within-firm relocation of jobs between establishments will have a minimal impact on the labor markets and would falsely be considered as job creation and job destruction in an establishment approach.

Identifying longitudinal firm histories and, by extension, the point of entry and exit, requires an operational definition of entrants and exits. In this paper, 'real' entrants are defined as firms that enter the market by starting new operations and creating new employment positions. Likewise, real exits correspond to firms that shut down and terminate all existing employment contracts. In between, firms are defined as continuing, also when they merge or split up activities. These definitions closely correspond to the concepts of entry and exit used in theoretical models of firm dynamics (Jovanovic 1982), as well as to the definitions that are implicitly assumed in most empirical studies (Caves 1998). Such *de novo* entry is opposed to entry by established firms which can take a variety of forms (Caves and Porter 1977). Likewise, exits by closing down activities differ from firms that transfer their activities to another legal entity. The definitions of entry and exit used in this paper are also in line with the ones recommended by OECD and Eurostat, as will be discussed below.

2.3.1 Linked employer-employee dataset

The register of Belgian employer firms is maintained by the Belgian National Social Security Office (NSSO) and is based on quarterly social security

[6] The unit of the 'firm' or more specifically the 'employer-enterprise' in the Belgian business register complies with the EU Regulation (EC) No 177/2008 for the harmonization of the national business registers for statistical purposes. It corresponds to "the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit." (Eurostat-OECD 2007)

declarations.[7] It covers all private firms with at least one employee in the period from 2003-2012, including 200 000 active firms and 2 500 000 employees on average per year.[8] Table 2.A.1 in the Appendix reports the number of firms and employees in the dataset, classified into eight industry groups.

The register is a linked employer-employee dataset. Both employers and employees are identified by means of a unique identification number. This information is exploited for the employee-flow method. The NSSO employer number is uniquely linked to the CBE number, which is the official firm identification number that is used by all government administrations. The CBE number enables us to implement results of the traditional record linking method, developed by Statistics Belgium, into the NSSO dataset. The CBE/NSSO number ensures good quality longitudinal firm records. Upon registration, new firms receive a CBE number which they keep for their entire lifetime. Unlike in many other countries, the firm identification number is unaltered in the case of a change of ownership or legal form.

A new CBE number is however assigned by the administration in a few situations: when a self-employed individual turns his or her business into a company, and when a firm changes ownership after bankruptcy. Furthermore, firm-induced ID changes occur for similar reasons as it is the case in other countries. For the purpose of accounting advantages or avoidance of liability, firms may exit by voluntary liquidation or bankruptcy and continue the same activities in a newly registered company.[9]

Firm identification numbers may also be created or disappear when legal entities are merged or split-up. Such changes in the firm structure may reflect actual mergers, acquisitions, break-ups or divestitures, but also mere administrative transfers of activities between legal units of the same controlling enterprise. A common example of the latter is the subdivision of the firm in smaller entities with separate firm identification numbers. For expanding firms, this is a

[7] The social security contributions are subject to strict control. The NSSO declarations are filled out electronically by the employer and missing declarations or unexpected changes in employment are checked by NSSO analysts. This ensures continuity of the firm identification number and make the data unlikely to be contaminated by measurement error.

[8] Temporary work agencies are left out from the analysis in this paper because the high job turnover in this industry confuses the discussion of average job reallocation.

[9] See for example Benedetto et al. (2007) for a discussion of these practices in the U.S.

way to remain below the size thresholds for legal obligations.[10] Other reasons why firms create additional ID numbers are tax advantages (the separate entities are not considered as part of the same firm) or limitation of liability. The practice is common under the form of enterprises controlling a network of local affiliates, but it is also used to distribute the firm activities across industries with differential regulations.

2.3.2 Firm linkage methods

Traditional record linking method

The traditional linkage method that is applied in this paper relies on probability-based matching and the use of supplementary data sources. It has been developed by Statistics Belgium within the Eurostat-OECD framework on business demography. Eurostat and OECD provide clear-cut definitions of enterprise 'births' and 'deaths'. Firm identification numbers that enter and exit for other reasons should be filtered out. A birth, for instance, "amounts to the creation of a combination of production factors with the restriction that no other enterprises are involved in the event" (Eurostat-OECD 2007, p. 34). Births should not include entrants due to restructurings of a set of enterprises such as mergers or break-ups, newly created enterprises after a change of legal form, take-overs of the activity of an existing enterprise, creations of additional legal units solely for the purpose of providing a single production factor or an ancillary activity, and so on. Likewise, a death "amounts to with the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event" (ibid., p. 51).

Statistics Belgium uses information from Commercial Court files and from the NSSO to identify changes in firm ID numbers and in firm structure. The Commercial Court provides information on official mergers, acquisitions and split-ups, and on changes in the CBE number. The link between the NSSO and the CBE number further help to track firms that change ID number.

The linkages are complemented with a probabilistic matching procedure. Similarities in name, address, and 4-digits industry code are used to compute probabilities that records refer to the same firm. Automatic and industry-specific ad-hoc rules are applied to verify the results. Although advanced software is

[10] In Belgium, small firms do not need to file full annual accounts or install a works council (with fewer than 100 employees, turnover below 7.3m EUR, and balance sheet total below 3.65m EUR).

adopted to minimize false (non-)matches[11], probability-based matching of firm records is subject to subjective evaluation and analyst intervention (Baldwin et al. 1992; Robertson et al. 1997). Matches based on partial identifiers are often imperfect and have to be checked manually. Moreover, major changes in discriminating identification numbers, e.g. name or telephone number, reduce the probability of a positive match, while such modifications often occur at the very moment a firm implements a legal or organizational change. Probabilistic matching is also less suitable for the identification of changes in firm structure such as mergers and split-ups. Therefore, the matching procedure is followed by extensive analyst review. All accepted matches, as well as an important part of rejected and probable matches are validated by making use of information on firm continuity and inter-firm linkages from a comprehensive business database that combines different administrative sources, business surveys, and statistical registers.[12]

Employee-flow method

The employee-flow method, also called 'labor-tracking' method, uses one main criterion to establish linkages between firm identification numbers: similarity of the workforce. Actual implementations of this method are basically similar in design. Changes in firm ID numbers and in firm structure are identified by tracing large clusters of employees that appear to 'move' from one firm identification number to another between two subsequent observations in time. The methods rely on the assumption that the simultaneous transition of a significant number of employees from one firm identification number to another is unlikely to be the result of individual worker mobility. Therefore, the actual linkage procedure generally starts from a minimum cluster of three to five employees, as for smaller clusters, there is a high probability that the employee flow merely represents individual job changes. This absolute threshold is supplemented with a set of

[11] The matching procedure used by Statistics Belgium is based on the Term Frequency – Inverse Document Frequency method.

[12] The validation process is carried out by making use of the comprehensive business 'datawarehouse' DBRIS. DBRIS is a relational database of all Belgian firms which links information at the firm level of a vast set of administrative sources (national register of legal entities, Annual Accounts, VAT declarations, Social Security declarations,…), business surveys (Structure of Enterprises Survey, Structure of Earnings Survey,…), and statistical registers on enterprise ownership and control structure (including consolidations, participations and FDI registered by the National Bank of Belgium and the European Group Register).

relative thresholds, which aim at avoiding false matches and at distinguishing between different types of firm restructurings. Due to the minimum cluster size, employee-flow methods are inappropriate for linking small firms. Yet they do achieve high coverage of linkages between larger firms, where sufficiently large clusters of employees can be followed over time.

The employee-flow linkages applied in this paper are generated by a simple linkage algorithm that consists of two stages. In a first stage, the set of pairwise ID numbers is identified that share a significant cluster of employees in two successive quarters. It includes all pairs for which at least five employees move from a first ID number in quarter q-1 (the 'predecessor') to a second ID number in quarter q (the 'successor').[13] The simultaneous transition of a significant number of employees in such a short time span is a first indication that the employee flow might be not the result of individual job changes. The second stage singles out the ID pairs that are retained as firm linkages. It consists of a set of decision rules that capture different forms of inter-firm relationships. The rules include thresholds for the relative cluster sizes, i.e. the size of the clustered employee flow relative to the total workforce of the firms involved. Section 2.B in the Appendix describes the full set of rules and their formal conditions. Three major rules cover 90 percent of all linkages and are briefly discussed below.

The first decision rule covers the major part of employee-flow linkages (57%) and captures links between ID numbers with largely identical workforces. Two firm identification numbers are linked if the employee-flow cluster represents at least 50 percent of the workforce of both the predecessor and the successor. This condition is a formal translation of our workforce-based definition of firm continuity: two successive firm identification numbers that employ mostly the same workforce, are considered to refer to the same firm.

Two other major rules identify links between smaller and larger firms. A firm may disappear from the dataset while continuing its activities as part of a larger entity. Such 'absorptions' by existing firms do not meet our definition of exit, and the transfer of workers to the merged entity does not correspond to the destruction and creation of jobs. To capture these events, a link is established if at least 75 percent of the workforce of an exiting firm is transferred to an already established firm. This second rule identifies 22 percent of additional linkages. The third rule captures the opposite case, when a significant part of the workforce of

[13] The minimum threshold of five employees is in line with other recent applications of the employee-flow method (Benedetto et al. 2007; Dixon and Rollin 2012).

a continuing enterprise is transferred to a newly created ID number. If the employee cluster coming from the established firm represents at least 75 percent of the workforce of the new entrant, a link between the two ID numbers is established. Such ‘split-offs’ cover an additional 11 percent of linkages. Mergers, break-ups and more complex forms of inter-firm linkages are identified with other decision rules described in the Appendix 2.B. They each cover only small parts of additional linkages.

The threshold values for the relative cluster sizes are to a certain extent arbitrary. One may be concerned that changing these values will have a significant impact on empirical estimates. However, the robustness checks presented in the Appendix B.2 show that they have not. The reason is that the employee clusters that link two firm ID numbers mostly represent close to 100 percent of the workforce of the predecessor, the successor, or both. Several robustness checks have been performed to test the sensitivity of the empirical results to the set of criteria imposed by the linkage algorithm. Relaxing or restricting the relative cluster size thresholds hardly affects the results. Reducing the set of decision rules has little impact either. This does not mean that improvements of the method could not be achieved, for example by deriving industry-specific thresholds from firms that do not change identification number.

Note that the employee-flow method uses a clear-cut operational definition of firm continuity that is directly translated into the linkage algorithm. If a significant part of one key input factor of the firm, the workforce, is moved from one administrative ID number to the next, it is defined as a continuing firm. This notion of firm continuity may or may not be appropriate for a particular research question, but at least it is unambiguous. The traditional method, by contrast, combines various notions of firm continuity as present in different supplementary data sources, derived from partial firm identifiers in the probabilistic matching procedure, and employed by the analysts in the validation process. The result is an ambiguous definition of firm continuity and it may be difficult, if not impossible, for the user of the dataset to ascertain whether it fits with the specific research question he wants to address.

Both linkage methods combined

We also construct longitudinal firm linkages that incorporate all information provided by both the traditional and the employee-flow method. Linkages edited in this way are the most accurate longitudinal firm records that can be obtained with the available methods. They will be used to calculate empirical benchmark measures, which serve as a reference to compare the results obtained by each of the two individual methods.

2.3.3 Re-estimating measures of entry, exit and growth

Measures of firm and employment dynamics presented in this paper are computed as year-by-year changes between June 30th of year t-1 and year t. The entry and exit of a firm are defined as the first and last year it reports positive employment. In between, firms are labeled as continuing. Continuing firms may have no employees in a given year.

Improved longitudinal linkages are first used to identify continuing firms which are misclassified as entrants and exits. They will be labeled as 'spurious' entrants and exits. As is the common practice, they are removed from the entry and exit populations to obtain improved measurements. Re-estimating firm-level growth measures is more challenging, as several firms can be interlinked in a given period. To our knowledge, no satisfactory solution has been suggested so far. We propose a simple solution for imputing employment growth at the firm level. Aggregate statistics then follow naturally from the revised firm-level observations.

The following example illustrates the problem. Suppose a link is identified between two firms that merge into a new administrative entity. The two firms, previously misclassified as exits, are now identified as continuing. The jobs of these firms are not lost, neither should the jobs that are transferred to the new entity be treated as job creation. But in the same period, job growth or decline may have occurred at the aggregate level of both firms, which does reflect actual job creation or destruction.

The approach adopted here is to first construct an aggregate event level including all firms interlinked in a given period from t-1 to t. Firm-level employment in t is then imputed by assuming the same growth rate for each firm involved in the event. The advantage of this imputation approach at the firm-level compared to solutions proposed elsewhere is that we do not change the firm counts and preserve the firm size distribution at the beginning of each period.[14]

[14] The approach of the U.S. Bureau of Labor Statistics (Pinkston and Spletzer 2002) and Statistics Canada (Dixon and Rollin 2012) is to collapsing both firms into a consolidated employer and calculate employment change at the level of this merged entity. For aggregate measures of job creation and destruction, our strategy yields the same result as this approach. The disadvantage of using a consolidated entity is that the firm counts will be

This allows for consistent estimations of firm-level measures that depend on firm size. The imputation procedure has a straightforward interpretation in the case of most types of events, as is discussed in Appendix 2.C.

Imputation of employment is performed on a year-by-year basis, i.e. for firms involved in an event in a given period t-1 to t. In the next period, we restart from registered employment in t and impute employment in t+1 for events in that period. Geurts and Van Biesebroeck (2014) have extended the imputation method over a five-year period. They found that firm ID numbers involved in a linkage event are more likely to be involved in another event in one of the following years than other firms. Reconstructing longitudinal employment histories of firms over several years thus quickly turns into a complex exercise in which multiple events of interlinked ID numbers have to be taken into account.

2.4 Characteristics of linked firms

The impact of missing linkages on empirical estimates crucially depends on the type of events that lead to missing firms linkages and, even more importantly, on the proportion of missing links by firm size class. Below, we discuss both elements before turning to the empirical results in Section 2.5.

2.4.1 Events leading to missing linkages

Table 2.1 summarizes the types of events that give rise to spurious entrants, spurious exits, and missing links of continuing firms. We use a basic typology of events that clarifies the main reasons behind missing linkages. The results are reported for each linkage method separately.

Spurious entrants mostly emerge from ID changes. An ID change is defined as a one-to-one link between two successive ID numbers of a firm that simply continues its operations with a new number. ID changes explain more than half of the spurious entrants identified by the traditional method (57%), and two thirds of the ones identified by the employee-flow method (65%). Another third of misclassified entrants, according to both methods, are due to split-offs of parts of incumbents or full break-ups of firms into new entities. The events may reflect an

inconsistent across time and, more importantly, the relation between firm size and firm growth will be biased. Indeed, the size of the consolidated entity is by construction larger than those of the original firms.

actual split-up of the firm, but also the creation of an additional legal unit within the same controlling enterprise.[15] Only few spurious entrants originate from mergers or more complex events that involve several firms. Spurious exits largely result from ID changes as well, as counterparts of spurious entrants. In addition, more than 40 percent of misclassified exits are explained by firms that are absorbed by an established firm or merged with other exits into a new firm ID number.[16]

Table 2.1 Types of events leading to spurious entrants, spurious exits and linked continuing firms

	Type of event (in percent of row total)					
a. Spurious entrants						
Linkage method:	ID change	Split-off or break-up	Merger	Combi-nation	Total	
Traditional method	57	36	1	6	100	*(n = 1 149)*
Employee-flow method	65	30	2	3	100	*(n = 952)*
Both methods combined	57	35	1	7	100	*(n = 1 867)*
b. Spurious exits						
Linkage method:	ID change	Absorption or merger	Break-up	Combi-nation	Total	
Traditional method	45	49	0	6	100	*(n = 1 469)*
Employee-flow method	54	42	1	3	100	*(n = 1 163)*
Both methods combined	49	43	1	7	100	*(n = 2 207)*
c. Linked continuing firms						
Linkage method:	Absorption	Split-off	All continuing	Combi-nattion	Total	
Traditional method	12	9	77	3	100	*(n = 8 630)*
Employee-flow method	46	31	13	11	100	*(n = 944)*
Both methods combined	14	11	71	4	100	*(n = 9 240)*

Note: Annual averages over the 2003-2012 period.

The number of continuing firms that are linked to another firm identification number strongly differs between the two linkage methods. The employee-flow

[15] Rapidly growing firms have an incentive to split-up activities into smaller legal units to remain below the size threshold for legal obligations.

[16] Comparable results are reported by Benedetto et al. (2007) and Hethey and Schmieder (2013). Using a more comprehensive typology of events, they also find that large shares of spurious entrants and exits are explained by ID-changes, split-offs and absorptions

method mainly traces incumbents that take over the workforce of an exiting ID number (the counterpart of a spurious exit due to absorption), or split off part of their activities into a new legal entity (the counterpart of a spurious entrant due to split-off). The traditional method identifies many more links, especially between two continuing firms. It predominantly captures links between large conglomerates of legal entities that are part of the same controlling enterprise, such as franchised businesses of a large company. The probabilistic matching procedure re-identifies these links in each successive period. However, as long as no restructuring of activities between the entities occurs, such links have little impact on the empirical dynamics measures, as will be shown below.

2.4.2 Spurious entrants and exits by size

Table 2.2 provides the main explanation for the biases in the empirical measures that will be discussed below. It presents the percentage shares of administrative entrants and exits that are identified as spurious ones by either of the linkage methods. The first column reports the shares in the total population of entrants or exits, the other columns give the shares in a given size class. The benchmark results based on both linkage methods combined show that one in ten administrative entrants and one in eight exits are identified as spurious. Both the traditional and the employee-flow method capture a much lower share of misclassified firms, which indicates a high degree of complementarity between the two methods.

Two important patterns emerge from Table 2.2. First, the probability that a new firm identification number corresponds to a spurious entrant increases dramatically with size, and the same holds for exits. This pattern is consistent across the two link methods. The traditional linkage method identifies about 5 percent of the smallest entrants and exits as being misclassified, which amounts to more than 40 percent in the largest size class. The employee-flow linkages reveal this pattern even more sharply. It shows that one in three entrants and exits with 5 to 9 employees and almost all entrants and exits with over 100 employees are brought about by ID changes or firm restructurings. The implication for the entry and exit measures is that missing links will have a larger effect on the employment shares of entrants and exits than on the firm entry and exit rates. The results in Table 2.2 highlight that newly registered firms with over 50 employees are most likely incumbents that continue operations – either in total or partially – with a new identification number. Likewise, if a large firm exits the dataset, there is a high probability that it refers to a continuing employer that has transferred its activities to another legal entity. The observation that firms entering the market with over 50 or 100 employees are exceptional is intuitive but rarely reflected in large-scale datasets. Research samples including a substantial amount of large entrants are a first indication that the longitudinal data may need more editing.

Table 2.2 Share of spurious entrants and exits by size

		Firm size (number of employees)					
	Total	1-4	5-9	10-19	20-49	50-99	100+
Entrants							
Unedited data *(n)*	*19 069*	*16 852*	*1 345*	*527*	*255*	*54*	*37*
Share of entrants identified as spurious (in percent of entrants in the unedited data)							
Traditional method	6	5	13	16	20	31	41
Employee-flow method	5	-	30	52	67	77	97
Both methods combined	10	5	36	55	70	80	97
Exits							
Unedited data *(n)*	*18 692*	*16 058*	*1 454*	*649*	*374*	*96*	*60*
Share of exits identified as spurious (in percent of exits in the unedited data)							
Traditional method	8	6	16	23	33	42	57
Employee-flow method	6	-	30	50	65	75	90
Both methods combined	12	6	36	54	69	78	91

Note: Annual averages over the 2003-2012 period.

The second observation from Table 2.2 is that the two linkage methods perform quite differently by firm size class. In all size classes above 5 employees, the employee-flow method captures two to three times more spurious entrants and exits than the traditional method, and it traces almost all misclassified firms that are identified when using both methods combined.[17] The added value of the traditional method in these size classes is rather low. The traditional method is however necessary for identifying misclassified firms in the smallest size class (1-4 employees), where employee-flow linkages are absent by construction. The close approximation between the employee-flow and benchmark linkages in medium and large size classes is a feature that will be reflected throughout the results discussed in the next section. For all empirical measures with an

[17] Benedetto et al. (2007) equally show for U.S. data that the employee-flow method has a nontrivial value added to other linkage methods for identifying missing links between firm identifiers.

employment component, the method performs better in avoiding bias than the traditional method.

It will be shown below that correctly identifying entry and exit has a dramatic impact on job reallocation measures, because the relatively small amount of misclassified firms represent important employment shares at entry and exit. An accurate distinction between what we have labelled as real versus spurious entrants and exits has however implications for firm-level analysis that reach far beyond the set of measures considered in this paper. Earlier studies that have made a similar distinction have found pronounced differences between the characteristics of the two types of entrants and exits. Treating them as a homogeneous group can lead to highly misleading conclusions about the empirical patterns of firm entry, exit and growth. One distinctive feature is size. Baldwin and Gorecki (1987) and Mata (1993), for example, have shown that entry by established firms and exits brought about by changes in the firm structure are many times larger than *de novo* entrants and exits by closure. The results in this paper confirm these findings. On average, spurious entrants and exits are eight times larger than real entrants and exits.[18] Other studies have demonstrated that entry and exit from ID changes, firm restructurings or diversifying firms also differ in characteristics other than size, such as in the determinants of entry (Acs and Audretsch 1989; Storey 1991), post-entry growth patterns (Mata et al. 1995; Geurts and Van Biesebroeck 2014), or the profitably and productivity at exit (Baldwin and Gorecki 1987).

2.4.3 Firms with imputed employment growth

Firm-level employment growth of linked firms is revised in each period from t-1 to t using the approach described in Section 2.3.3. The revision applies to spurious exits reclassified as continuing firms and to continuing firms that are linked to another ID number. A moderate share of firms are affected and, although substantial, the impact on the empirical measures will be less important than the impact of filtering out spurious entrants and exits.

Table 2.A.3 in the Appendix presents the share of firms of which employment growth is considerably revised after the imputation procedure. We only report firms that subject to a relative adjustment of more than 10 percent compared to the unedited data. The benchmark results based on both link methods combined

[18] The average sizes of real versus spurious entrants and exits are reported in Table 2.A.2 in the Appendix.

show that on average 3.8 percent of all active firms in a given period are concerned. Similar to spurious entrants and exits, this share increases with size and amounts to more than 7 percent of the firms with over 50 employees. The traditional link method affects more firms than the employee-flow method, especially in size classes under 50 employees. The employee-flow method, however, has a greater impact on larger firms, and will lead to a more substantial revision of total job reallocation measures.

2.5 Results

This section discusses the sensitivity of empirical measures of firm and employment dynamics to missing links in longitudinal firm histories. The measures are evaluated before and after implementing each linkage method, and compared with benchmark results based on both methods combined. All measures are computed as year-by-year changes between June 30th of year *t*-1 and year *t*. Along with the description of our findings, we also briefly address the following three issues: the comparison with results for other countries; the implications for comparative statistics; and the policy implications. The results in this section refer to the total private sector. As shown in the Appendix 2.F, the results are qualitatively the same at the disaggregated industry level, but the biases are exacerbated in industries with a relatively high share of large firms, while they are smaller in industries with predominantly small firms.

2.5.1 Entry and exit dynamics

In the previous section, it has been shown that the share of misclassified firms strongly increases with firm size, and that spurious entrants and exits are, on average, much larger than firms actually entering and exiting the market. The implication for the empirical measures is that missing links will have a larger effect on job flows by entry and exit than on the firm turnover rates.

Table 2.3 indeed shows that entry and exit rates are only moderately revised downwards after applying the linkage procedures. The results in the first row are based on the unedited administrative data. The next rows show the revised measures after spurious entrants and exits have been filtered out. The lower panel of the table shows the percent bias in the measures compared to the benchmark results. Each linkage method produces entry and exit rates that are slightly below 10 percent, which is in line with results for other European countries (Reynolds et al. 1994; OECD 2015). The traditional method yields entry and exits rates that most closely correspond to the benchmark results since it performs well in capturing misclassified firms in the smallest size classes, where the majority of entrants and exits is located and employee-flow links are absent by construction.

Table 2.3 Summary statistics of entry and exit

	Entry measures			Exit measures		
	Entry rate (%)	Job creation rate (%)	Average size (empl.)	Exit rate (%)	Job destruction rate (%)	Average size (empl.)
a. By linkage method						
Unedited data	9.6	2.5	3.3	9.4	3.1	4.1
Traditional method	8.8	2.1	2.9	8.6	2.2	3.1
Employee-flow method	9.1	1.5	2.0	8.8	1.6	2.3
Both methods combined	8.7	1.4	2.0	8.3	1.5	2.3
b. Percent bias vs. both methods combined						
Unedited data	*11%*	*81%*	*64%*	*13%*	*102%*	*78%*
Traditional method	*2%*	*50%*	*46%*	*4%*	*48%*	*35%*
Employee-flow method	*5%*	*7%*	*1%*	*6%*	*8%*	*2%*

Note: Annual averages over the 2003-2012 period. The entry (exit) rate represents the share of entrants (exits) in all active employers of a given period. The job creation (job destruction) rate represents the employment share of entrants (exits) in total employment of a given period.

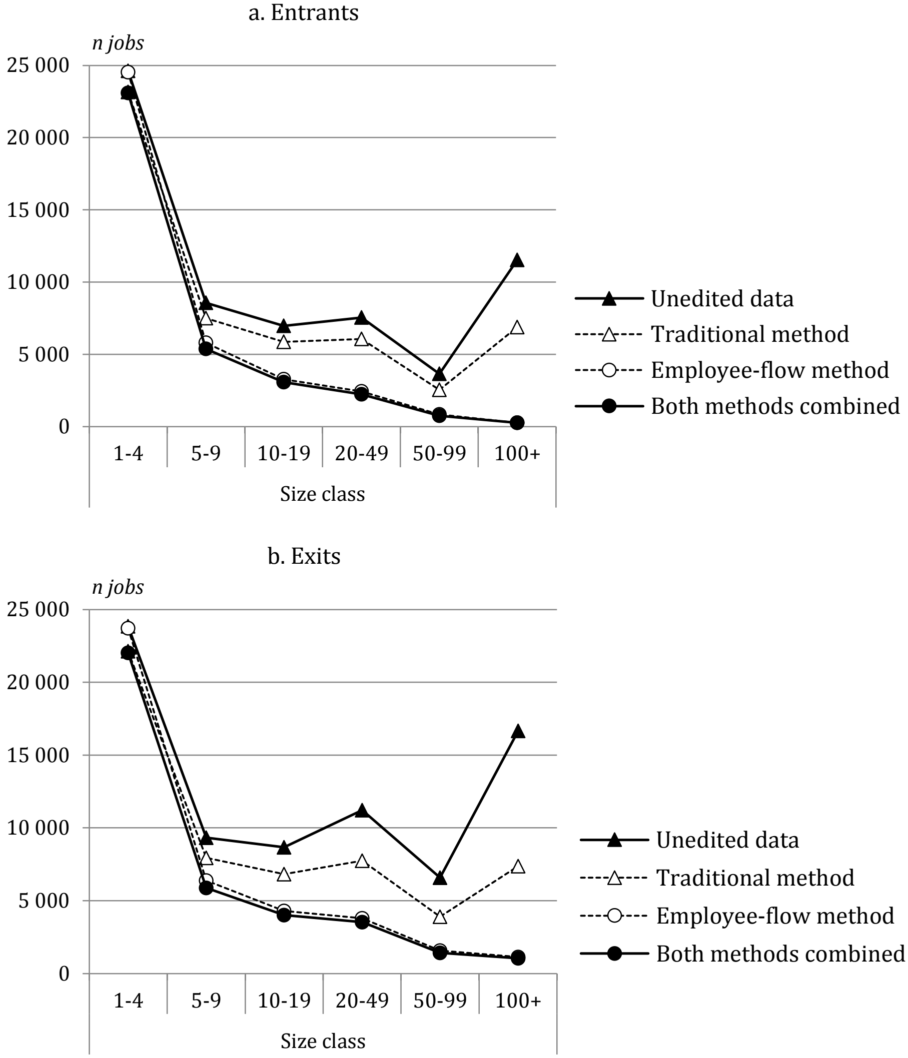
In contrast to firm turnover rates, employment measures at entry and exit are considerably revised downwards after spurious entrants and exits are filtered out. When based on the unedited data, the job creation and job destruction rates are overestimated by 81 percent and 102 percent respectively, and average entry and exit sizes by 64 percent and 78 percent respectively. The employee-flow method strongly reduces these biases and produces results close to the ones that are obtained by both linkage methods combined. The traditional method corrects the initial biases only by about half. The large number of additional links that this method identifies in the smallest size class (see Table 2.2) account for only small shares of aggregate job reallocation and contribute little to bias reduction.

The quantitative impacts of the linkage methods reported here are obviously country-specific, since they depend on the incidence of missing links in the raw dataset. Knowing that the Belgian administrative firm ID is relatively consistent,[19] the large downward revision of the job reallocation rates after correcting for missing linkages is all the more striking. Comparable statistics for other countries are scarce, but data from available studies show a similar large impact on job flow measurements at entry and exit. Persson (1998), Korkeamäki and Kyyrä (2000), and Hethey and Schmieder (2013), using employee-flow linkages to correct Swedish, Finnish, and German data respectively, equally find that job reallocation at entry and exit is reduced by about 50 percent when misclassified firms are filtered out. From a policy perspective, the revised outcomes can have significant implications, as will be discussed below.

The employment distributions at entry and exit shed more light on the above results. The upper panel of Figure 2.1 presents the distribution of total employment created by new firms at entry, and the lower panel shows the employment distribution of firms in the year of exit. The top lines represent results based on unedited data, while the other three lines show the results after applying the linkage procedures. Missing linkages strongly shift the distributions to the right. The unedited data falsely suggest that an important amount of jobs is created by medium and large entrants, and likewise that more than half of job loss due to exit is brought about by medium and large firms exiting the market. These patterns are only partly corrected by the traditional method. The method fails to identify an important part of spurious entrants and exits in larger size classes and leaves a considerable upward bias in the middle and right tale of the distributions. The employee-flow method, by contrast, being more effective in tracing missing linkages of larger firms, strongly reduces the initial biases. Results obtained by this method reveal that job creation by entrants and job destruction by exits is almost entirely concentrated in the smallest size classes. Moreover, this method yields strongly right-skewed distributions that closely correspond to the ones obtained by using both methods combined.

[19] As mentioned above, Belgian firm ID numbers do not change after a change of ownership or legal form, in contrast to many other countries.

Figure 2.1 Employment distribution of entrants and exits



Note: Annual averages over the 2003-2012 period.

Many studies have found that the firm size distribution at entry is right-skewed and that the likelihood of a firm's exit declines with its size (Cabral and Mata 2003; Caves 1998). Even a small number of larger entrants or exits may, however, represent important employment shares at entry or exit. This is indeed the pattern that is suggested by the unedited data and the traditional method. Improved longitudinal data instead reveal that small firms do not only represent the major part of units but also the major part of employment at entry and exit. The benchmark results show that firms that start with less than 10 employees represent more than 80 percent of total employment at entry, and likewise that small firms account for the bulk of job destruction due to exit. Firms with more than 50 employees barely contribute to job flows at entry or exit. In summary, improved linkages reveal that the firm size distributions at entry and exit are confined to a narrow range of small size classes. Although this is an intuitive result for entrants, it is rarely observed in research samples based on administrative datasets. Studies that use a sample with a broad range of firm sizes already at entry must be investigating a population that covers also other firms than genuine start-ups. As discussed above, failing to properly distinguish between real and spurious entrants has important implications for the analysis of entry characteristics and post-entry patterns.

Given their policy relevance, statistics on job creation by start-ups and job destruction by firm exits are among the standard indicators produced by official institutes. One example are the comparative statistics published by Eurostat and OECD.[20] Figure 2.A.1 in the Appendix shows cross-country comparisons from these sources of the job creation rate by entrants, and the employment share of large firms at entry – both in the manufacturing sector. Looking at the results, one is puzzled by the remarkable divergence between European countries. The job creation rate of start-ups in France, for example, is reported to be ten times higher than in Germany. Consistent with our discussion above, France also reports an extremely high share of large firms in the population of entrants. We have also included results for Belgium based on our edited versions of the dataset. Interestingly, the traditional method ranks Belgium among the countries with middle to high values for both statistics, while the employee-flow method ranks it among the ones with the lowest values. It is not clear how the national datasets have been edited, but the analysis in this paper suggests that the large country differences may be explained by the quality of the longitudinal firm linkages. The results raise questions about the use of traditional linkage methods, recommended OECD-Eurostat, for obtaining internationally comparable results. As argued above, the comparability could be strongly increased by using an employee-flow method to edit the national datasets. The method is based on a linkage algorithm that only depends on a set of observable characteristics of firm

[20] The OECD and Eurostat statistics discussed here are derived from harmonized national business registers. Countries can either apply their own method to obtain consistent longitudinal data or follow the traditional linkage approach recommended by Eurostat-OECD as discussed above.

continuity, and can be standardized across countries. One requirement is the availability of linked employer-employee data.

2.5.2 Size profile of firm growth

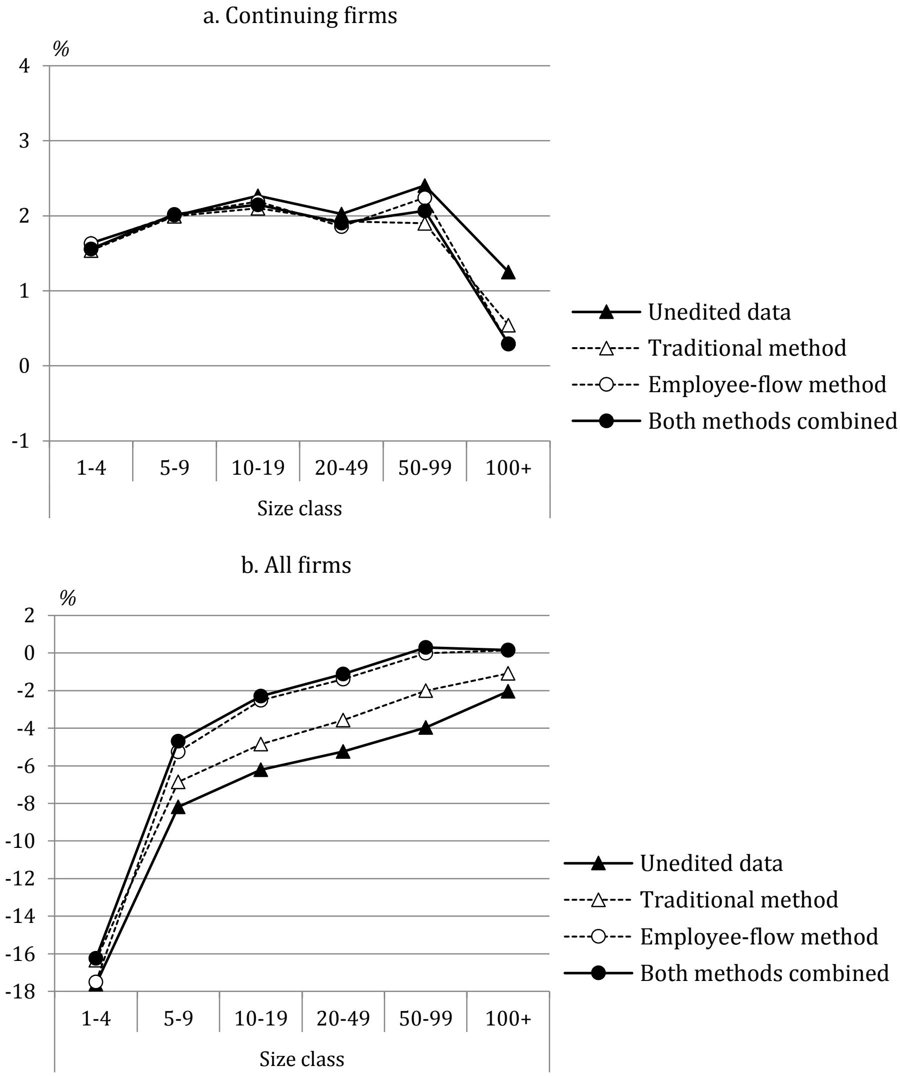
As explained in Section 2.3.3, firm-level growth rates are revised when exits are reclassified as continuing firms, and when continuing firms are linked to another firm identifier. A simple approach has been applied to impute employment of these firms. This section presents an example of the sensitivity of firm-level growth estimates to missing longitudinal linkages. Is shows estimations of the relationship between firm size and the firm growth rate, both in the sample of all firms, and in the sample of continuing firms only. The regression model and estimation method are described in Appendix 2.D.[21] The coefficient estimates with standard errors are reported in Table 2.A.4.

From the previous results, it can be expected that growth rates of large firms will be more sensitive to longitudinal linkage errors than the ones of small firms. Because missing linkages strongly affect the exit population in particular, it can also be expected that growth rate estimates of all firms will be affected more severely than those of continuing firms only.

The upper panel of Figure 2.2 plots the size coefficients of the regressions for firms that continue between t-1 and t. The point estimates represent the mean employment growth rates of a given size class of firms, which are the net result of job creation by expanding firms and job destruction by contracting firms. As expected, results for larger firms are more sensitive to missing links between identification numbers. The average growth rate of firms with over 100 employees is 1.3 percent per year when based on the unedited data, but revised downwards to 0.3 percent in the benchmark results. Revisions based on the traditional and employee-flow method are quite similar. At least as important for firm-level analysis is the increase in the precision of the estimates that is reached after the data are edited (Table 2.A.4 in the Appendix). Improved longitudinal linkages eliminate large administrative leaps in the employment histories of individual firms which leads to a reduction in the standard deviations of the growth rates.

[21] Regressions of continuing firms include 1.6 million firm-year observations; those of all firms 1.7 million.

Figure 2.2 Firm-level growth rates (in percent) of continuing firms and all firms by size



Note: Annual averages over the 2003-2012 period.

The lower panel of Figure 2.2 reports the growth estimates for all active firms in *t*-1, i.e. including the ones that exit in *t*. The plotted curves show a positive relationship between firm size and firm growth, which is in line with Haltiwanger et al. (2013) who use a similar estimation method. The difference between the slopes of the upper and lower panels is explained by the higher exit rates of firms in smaller size classes. The lower panel shows that growth rates of all firms in all size classes are considerably revised upwards after the data are edited by the linkages procedures. This means that reclassifying spurious exits as continuing firms overrules the downward revision in growth rates of continuing firms. The revision is most substantial in size classes between 5 and 100 employees, where the benchmark growth rates are 4 percentage points higher than the ones obtained with the unedited data. A considerable gain in the precision of the estimates is reached as well, mainly because the minimum growth rates of firms misclassified as exits are replaced with values closer to the mean (Table 2.A.4). Standard deviations of the benchmark estimates are about one-fourth smaller than the ones based on the unedited data. The employee-flow method proves to be more successful in reducing bias in the growth estimates in size classes above 5 employees: both the mean and standard deviations are close to the benchmark results. In line with the above findings, this is explained by the greater capacity of the method to capture spurious exits in medium and large size classes. The traditional method, which misclassifies an important share of these firms, reduces the biases by only half.

2.5.3 Job creation and job destruction

The results up till now suggest that a poor strategy to identify real entrants and exits, and true employment gains and losses when firms are merged or split-up will strongly affect total job reallocation rates. The reason is that misclassifications increase with firm size, and thus account for important shares of aggregate job flows. To document aggregate measures of job creation and destruction, we follow Davis et al. (1996a), and decompose the net employment growth rate into the job creation rates by entry and by expansion, and the job destruction rates by exit and by contraction.[22]

Total job creation and job destruction rates

Table 2.4 shows that missing linkages introduce a considerable upward bias in each of the four job reallocation rates. Average annual employment growth was 1.03 percent in the period of observation (2003-2012).[23] Following the benchmark method, the net employment growth was the result of an average annual job creation rate of 7.06 percent and a job destruction rate of 6.03 percent. When based on the unedited data, these measures are overestimated by 28 and

[22] The decomposition is given in Appendix 2.E.

[23] By definition, net employment growth is not affected by the linkage procedures since they only redefine the reallocation of jobs across firms.

32 percent, respectively. The biases are mainly due to the overestimations of the job reallocation rates by entry and exit discussed before. The relative biases in the job reallocation rates of expanding and contracting firms are much smaller, but further add to the overestimation of total job flows. In line with the previous results, Table 2.4 confirms that the employee-flow method performs very well in reducing the initial biases and yields job reallocation rates that are close to the benchmark results. The traditional method leaves a substantial upward bias in each of the measures.

Improving longitudinal firm linkages not only reduces the total supposed amount of job flows on the labor market, but also affects the magnitude of the contribution of different types of firms to net employment creation. This may have significant policy implications as more accurate measurements may lead to different conclusions about the importance of specific groups of firms for job creation. We illustrate this with two examples: the contribution of new versus established firms to employment growth in a given period, and the contribution of small versus large firms. The findings regarding the first issue represent a general result that is obtained after improving firm linkages; the results we report on the second issue are country specific and may not necessarily be observed in other datasets.

Support measures for start-ups are often motivated by statistical evidence about the huge amount of new jobs they create every year. Spurious entrants in the data, however, make their contribution to employment look much larger than it is. Table 2.4 shows that new firms contribute much less to job creation while established firms destroy less jobs than inadequately edited data suggest. The benchmark results reveal that jobs created by new firms represent a mere 1.39 percent of total employment in a given year (column 3); while the traditional method and the unedited data overestimate this share by 50 to 81 percent. On the other hand, job destruction by incumbent firms (last column) proves to be much smaller if based on good quality longitudinal data. Net annual employment growth of established firms is not 1.05 percent, as suggested by the traditional method, but only 0.36 percent. The bias reduction in these measures obtained by improving firm linkages will be qualitatively the same in other datasets; fragmented evidence for other countries provided in Section 2.5.1 even suggests a remarkable quantitative similarity with the results in this paper.

Table 2.4 Annual job creation and job destruction rates

	Net growth rate	Job creation rate			Job destruction rate			Net growth rate of establ. firms
		Total	By entry	By expansion	Total	By exit	By contraction	
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
a. By linkage method								
Unedited data	1.03	9.01	2.52	6.48	7.98	3.06	4.92	-1.49
Traditional method	1.03	8.17	2.09	6.08	7.14	2.24	4.89	-1.05
Employee-flow method	1.03	7.24	1.49	5.75	6.21	1.64	4.57	-0.46
Both methods combined	1.03	7.06	1.39	5.67	6.03	1.52	4.51	-0.36
b. Percent bias vs. both methods combined								
Unedited data		*28%*	*81%*	*14%*	*32%*	*102%*	*9%*	
Traditional method		*16%*	*50%*	*7%*	*18%*	*48%*	*9%*	
Employee-flow method		*3%*	*7%*	*1%*	*3%*	*8%*	*1%*	

Note: Annual averages over the 2003-2012 period. The job creation (job destruction) rate represents the employment share of a given class of firms in total employment of a given period. The net growth rate of established firms equals the job creation rate by expansion minus the job destruction rates by exit and contraction.

Missing firm linkages also considerably affect the allocation of employment growth across different firm size classes. Figure 2.2 above already showed that correctly classifying firms and their individual employment histories leads to an upward revision of firm-level growth rates. The growth rates by firm size class were estimated conditional on industry and year dummies. Figure 2.A.2 in the Appendix presents the impact on net employment growth aggregated over all established firms in the total private sector. Improved linkages yield a dramatic change in the contribution of different size classes to net job creation. While the unedited data and the traditional method suggest that net job creation of established firms in all size classes is negative, the employee-flow and benchmark method lead to an upward correction that is increasing in firm size. Both methods reveal that large firms actually contribute positively to employment growth.

In summary, results based on unedited data or on the tradition method overestimate the contribution of entrants and small firms to job creation, lending support to the common perception that these groups of firms contribute disproportionally to net employment growth. Results based on carefully edited longitudinal data challenge this view. Instead they indicate that job creation by entrants actually very low, and that large established firms contribute a great deal more to employment growth in the Belgian economy than smaller firms.

Annual variation in job creation and destruction

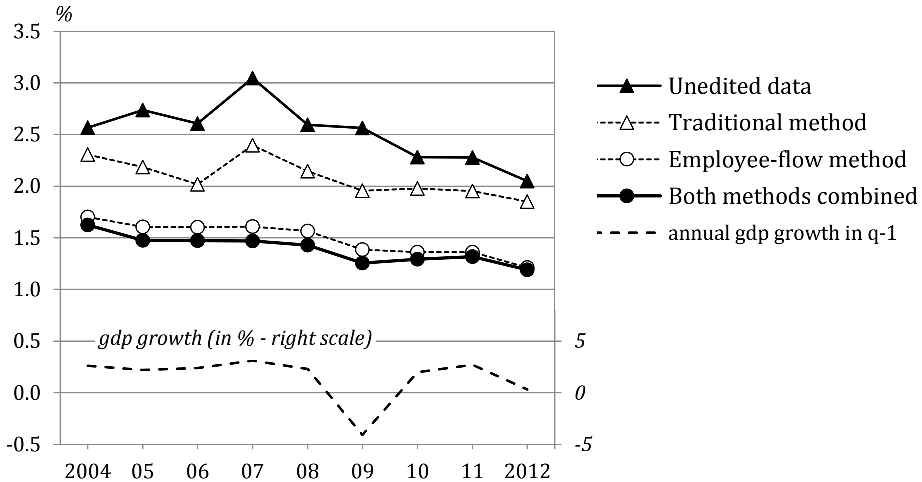
To conclude this overview of empirical measures, we discuss the impact of missing linkages on time series of aggregate job flows. The bias in the job reallocation rates is found to vary significantly over time. This is explained by annual fluctuations in the number of medium and large firms that change identification number or firm structure. The job reallocation rates by entry and exit most clearly illustrate the problem.

Figure 2.3 shows that annual variation in the job creation rate by entry and the job destruction rate by exit is strongly reduced when spurious entrants and exits are filtered out. The unedited data and the traditional method report large annual fluctuations, while the employee-flow and benchmark results reveal that both job creation by entry and job destruction by exit are rather non-volatile, with the largest year-to-year change corresponding to the recession period of 2009. One typical pattern is explained by ID changes: fluctuations in the employment shares of firms that change ID number lead to symmetric increases and decreases in the job creation rate by entry and the job destruction rate by exit. This effect is most noticeable in the results for 2007: the unedited data report a leap in both the entry and the exit rates, which is entirely absent in the benchmark results. Other studies have reported a similar smoothing of time series of entry and exit after applying improved longitudinal linkages (Jarmin and Miranda 2002; Hethey and Schmieder 2013).

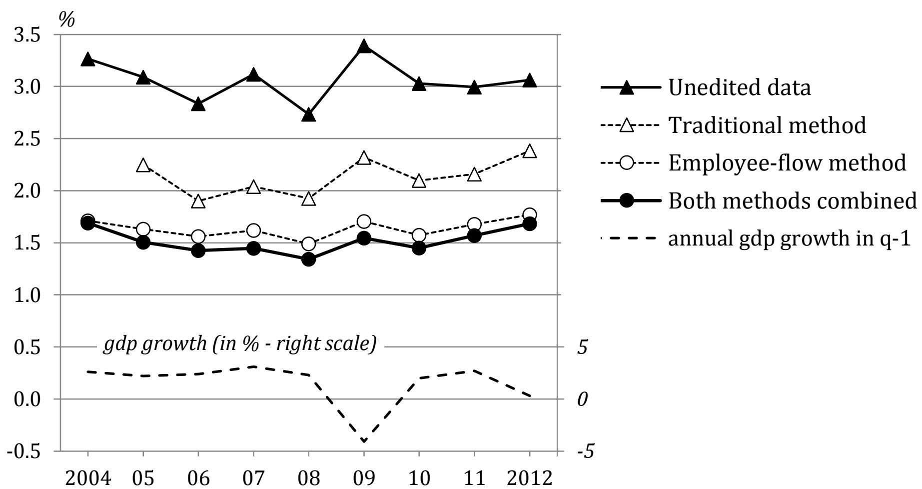
Is this stable pattern of job reallocation from firm entry and exit more plausible than the annual fluctuations suggested by the unedited data and the traditional method? The correlation of the entry rate with GDP growth suggests it is. Business formation is considered to be procyclical, and especially job creation by entry is found to covary positively with output growth (Campbell 1988). The employee-flow results strongly reflect this feature; annual changes in the job creation rate by entry show a high positive correlation with GDP growth of the previous quarter (0.86), which adds support to the reliability of this identification strategy for entry and exit. The correlation is only half as large (0.42) for entry rates based on the traditional method. Misclassifying continuing firms as entrants thus introduces substantial spurious variation in job creation by entry over time, which weakens the correlation with the business cycle. ID changes, mergers and split-ups are indeed mainly driven by legal, tax or administrative motivations and less by macro-economic fluctuations.

Figure 2.3 Annual job reallocation rates by entry and exit

a. Job creation rate by entry



b. Job destruction rate by exit



Note: The job creation (job destruction) rate represents the employment share of entrants (exits) in total employment.

2.6 Conclusion

We have discussed in this paper a series of commonly-used empirical measures on firm and employment dynamics and have exemplified to what extent they are biased due to missing firm linkages in the underlying data. The study has focused on firm-level measurements, but the problems and solutions are similar for establishment analysis based on administrative data. Measures with an employment component are found to be most seriously biased, whether it be firm-level growth estimates or aggregate job flow measures. The most important source of bias is the spurious identification of firm entry and exit. Medium and large firms misclassified as entrants or exits distort the size distributions at entry and exit and lead to large overestimations of job reallocation in the economy. Missing firm linkages also affect the characteristics of successful firms and lead to an underestimation of firm-level growth rates, especially for larger firms. The biases are exacerbated in time series and at the more disaggregated industry level.

Two prevailing solutions for improving longitudinal linkages have been evaluated: a traditional linkage method and an employee-flow method. In addition, we have proposed a consistent approach to impute employment histories of linked firms such that the edited data can be used for firm-level analysis. The two linkage methods are clearly complementary. The traditional method is preferable for the study of firms in the smallest size class (1-4 employees), where employee-flow linkages are absent by construction. The employee-flow method, however, performs much better for improving firm linkages in medium and large size classes. It is shown that this method is more suitable for obtaining reliable estimates of aggregate job reallocation and of firm-level measurements that depend on firm size.

An additional advantage of the employee-flow method lies in its use for international comparability. The method uses an economic notion of firm continuity, defining it as continuity in one of the firm's key production factors, the stock of employees. This definition is translated in linkage algorithms that use general criteria and are, unlike the traditional method, independent of country-specific data characteristics. Using employee-flow methods to harmonize longitudinal business databases for research could not only produce more reliable but also more comparable results across countries on firm and employment dynamics. One example highlighted in this paper are comparative statistics on business demography published by Eurostat and OECD. The analysis suggests that large country differences may be explained by different quality of longitudinal linkages in the national datasets, and that using a standardized employee-flow method could strongly increase comparability.

Improving longitudinal firm linkages in large-scale datasets has implications for firm-level analysis that reach beyond the set of measures discussed in this paper. Our findings suggest that existing empirical evidence of related studies on firm dynamics and entrepreneurship may be reconsidered.

A first area is the study of start-ups and post-entry performance. Consistent firm linkages enable the researcher to clearly distinguish between firms that enter the market *de novo* and entry by established firms - for example following an ownership change, a restructuring or a merger. Several authors have highlighted that the two types of entry fundamentally differ (Dunne et al. 1988; Geroski 1995; Mata et al. 1995). Firms that re-enter after a control change already have a better idea of their own productivity, tend to be larger, are less likely to fail, and exhibit less dynamic growth patterns. Their features may override the distinct patterns of genuine start-ups if both populations are confounded. As an example, we have shown that the employment distribution at entry is reduced to a narrow range of small size classes when we focus on *de novo* entrants. This intuitive feature is rarely observed in large-scale samples. Building on the same dataset, Geurts and Van Biesebroeck (2014) show that firm growth rates are strongly increasing in firm size in the first years after entry, but that this pattern is obscured or even reversed when the administrative sample is taken at face value.

A second area that may gain from improved longitudinal data is research on the determinants of exit and the performance of successful firms. Similar to entrants, empirical evidence has shown that real exits and exits brought about by take-overs, mergers or split-ups strongly differ in size, profitably and productivity (Baldwin and Gorecki 1987; Caves 1998). By identifying firm linkages, one can avoid confusion between such restructuring events and real economic exits. One implication we have highlighted is that misclassifying surviving firms as exits leads to an underestimation of the performance of successful firms, and biases the size distribution at exit. Moreover, missing firm linkages introduce measurement error and outliers in firm-level estimates. A significant gain in precision can be achieved when longitudinal firm histories accurately observed.

Being capable of tracing firm linkages over time has other applications in the field of firm dynamics that this paper has only hinted at. Ownership changes, mergers and acquisitions, break-ups and spin-offs are changes in the control structure of firms that researchers often do not want to abstract from, as we have done in this paper, but to study as events of economic importance. The employee-flow method in particular has been proved to be a powerful tool for identifying changes in the firm structure over time. A few studies have already used this method to investigate spin-offs (Eriksson and Kuhn 2006; Muendler et al. 2012), mergers and acquisitions (Mikkelson et al. 2006; Pesola 2009), and other forms of inter-firm relationships Benedetto et al. (2007). Many more applications are still to be explored.

Appendix

2.A Tables and figures

Table 2.A. 1 Eight main industries and NACE Rev.2 classes

Nace Rev. 2 classes	Employer firms (n)	Employees (n)	Average firm size (employees)
1. Agriculture Section A	4 097	20 912	5.1
2. Manufacturing and energy Section B, C, D, E	19 312	567 450	29.4
3. Construction Section F	26 354	207 510	7.9
4. Wholesale and retail trade Section G	55 208	480 235	8.7
5. Accommodation and food services Section I	20 524	120 042	5.8
6. Business support services 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, hedge funds 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; Rental and leasing: Nace 77.1, 77.3, 77.4; Security: Nace 80; Services to buildings: Nace 81; Administrative services: 82; Repair of ICT: Nace 95.1	34 000	388 616	11.4
7. Mixed business & household services 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: Nace 58; Movies, radio and television: Nace 59, 60; Telecommunication: Nace 61; Banks, credit and insurances institutions: Nace 64.190, 64.921, 64.922, 64.92, 65; Real estate: Nace 68; 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; Creative, arts and entertainment: Nace 90; Sports and recreation: Nace 93; Repair of household goods: Nace 95.2; Personal service activities: Nace 96	28 416	346 260	12.2
8. Human health and social work Section Q	11 286	378 163	33.5
Total	199 197	2 509 188	12.6

Note: Annual averages over the 2003-2012 period. Industries that are not in the listed categories are excluded from the analysis, i.e. primarily public sector organizations and temporary work agencies.

Table 2.A. 2 Average size (number of employees) of real versus spurious entrants and exits

	Real entrants	Spurious entrants	Real exits	Spurious exits
Linkage method				
Unedited data	3.3	-	4.1	-
Traditional method	2.9	10.1	3.1	14.1
Employee-flow method	2.0	27.1	2.3	30.5
Both methods combined	2.0	15.5	2.3	18.1

Note: Annual averages over the 2003-2012 period.

Table 2.A. 3 Share (in percent) of active firms in *t*-1 with imputed employment in *t*

		Firm size (number of employees) in *t-1*					
	Total	1-4	5-9	10-19	20-49	50-99	100+
Linkage method							
Traditional method	3.1	2.8	3.5	3.8	4.2	4.5	4.1
Employee-flow method	1.1	0.1	1.8	2.9	4.0	5.3	6.2
Both methods combined	3.8	2.8	4.7	5.6	6.5	7.3	7.3

Note: Annual averages over the 2003-2012 period.

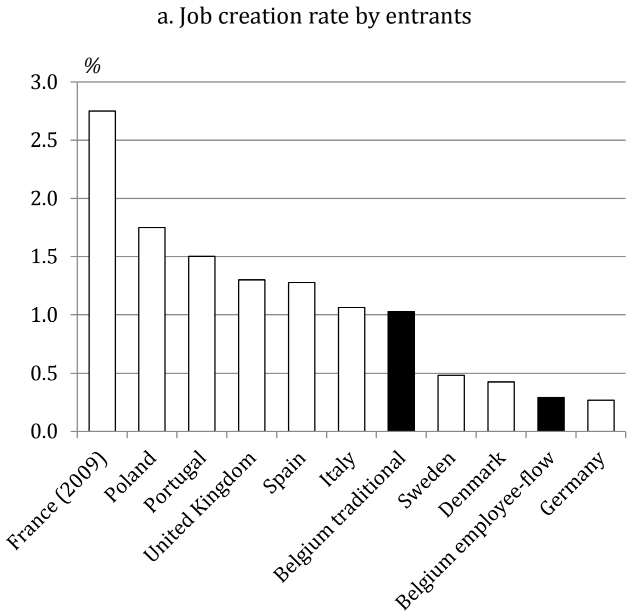
To clarify the impact of revising employment of linked firms on the empirical measures, this Table only reports the share of firms which are considerably affected by the imputation procedure, i.e. for which employment in *t* is revised by more than 10 percent. The revision is smaller for most of the linked firms. The main reason is that the traditional method establishes many links between continuing firms that do not relate to ID changes or firm restructurings and have little impact on employment reallocation. Another reason is that large firms are often linked to very small ones. If a large firm creates a small additional legal unit or absorbs a small entity, the revision of its employment will be relative small.

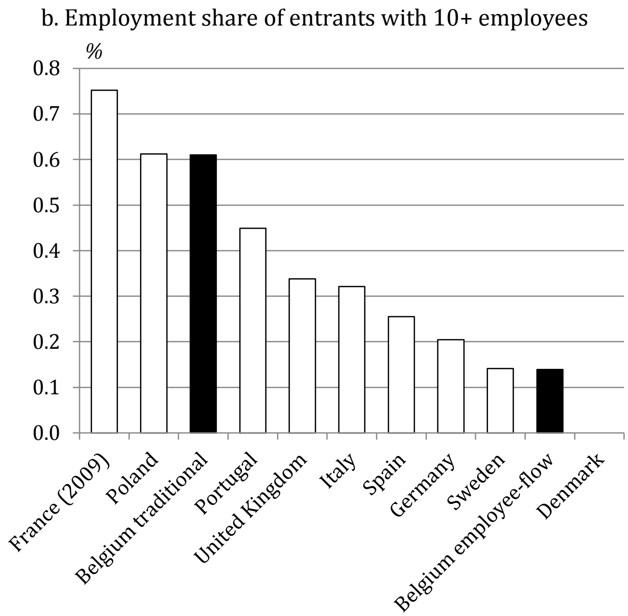
Table 2.A. 4 Firm-level growth estimates of continuing and all firms by linkage method

	Size class (number of employees)					
	1-4	5-9	10-19	20-49	50-99	100+
a. Continuing firms						
	Growth rate (%)					
Unedited data	1.63	2.00	2.27	2.02	2.40	1.25
Traditional method	1.54	1.99	2.10	1.93	1.90	0.54
Employee-flow method	1.63	2.01	2.19	1.86	2.24	0.29
Both methods combined	1.56	2.02	2.15	1.90	2.07	0.30
	Standard deviation					
Unedited data	.077	.085	.078	.075	.120	.098
Traditional method	.073	.081	.074	.072	.114	.093
Employee-flow method	.068	.074	.068	.066	.106	.086
Both methods combined	.067	.073	.068	.065	.104	.085
b. All firms						
	Growth rate (%)					
Unedited data	-17.62	-8.19	-6.21	-5.24	-3.95	-2.04
Traditional method	-16.35	-6.85	-4.84	-3.57	-2.00	-1.08
Employee-flow method	-17.51	-5.26	-2.51	-1.39	-0.01	0.14
Both methods combined	-16.24	-4.69	-2.29	-1.11	0.29	0.17
	Standard deviation					
Unedited data	.129	.145	.135	.130	.208	.172
Traditional method	.114	.129	.119	.115	.185	.153
Employee-flow method	.100	.112	.104	.100	.161	.133
Both methods combined	.098	.110	.102	.098	.158	.129

Note: Annual averages over the 2003-2012 period.

Figure 2.A. 1 Job creation rate by entrants in Manufacturing and Energy in European countries (2012)

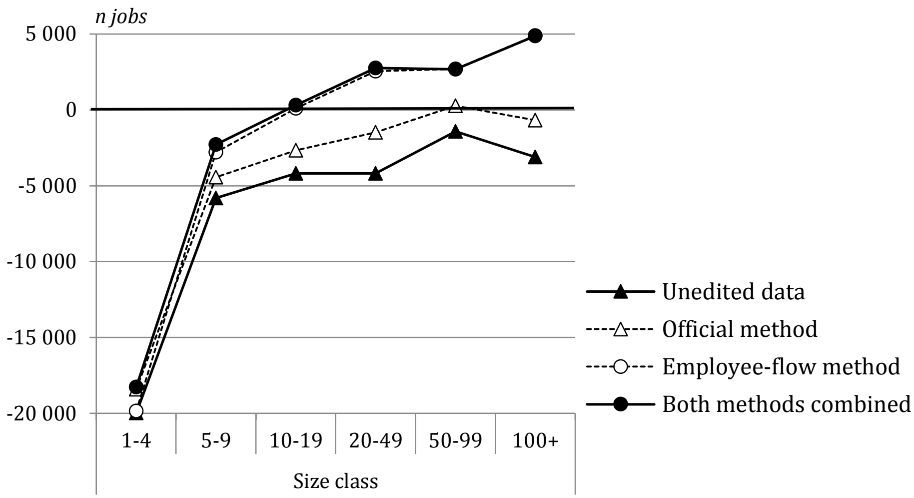




Note: Results for 2012 unless otherwise stated. Panel b shows the employment share of firms with 10 employees or more in total employment at entry.

Sources: Panel a: OECD (2015); Panel b: Eurostat Structural Business Statistics available online at http://ec.europa.eu/eurostat/web/structural-business-statistics. Results for Belgium are based on the data used in this paper.

Figure 2.A. 2 Net job creation of established firms by size class



Note: Annual averages over the 2003-2012 period.

Net job creation is calculated as the sum of job creation by expanding firms minus job destruction by exits and contracting firms. Firms are classified into the size class based on the average number of employees in *t*-1 and *t.* For a discussion see Davis et al. (1996b). The figure shows net employment growth aggregated over all established firms in the total private sector and therefore differs from the firm-level estimates in Figure 2.2 in Section 2.5.2, which represent firm-level growth rates conditional on industry dummies.

2.B Employee-flow linkage algorithm

Decision rules

An employee-flow link between two different firm identification numbers is established if a cluster of at least 5 employees moves from one identification number in quarter q-1 (the 'predecessor') to another identification number in quarter q (the 'successor'), and if the decision rules presented in Table 2.A.5 are met. The first three types of linkages represent the major share of links (90%).

Table 2.A. 5 Type of employee-flow linkages by decision rules

	Decision rules					
	Number of predecessors to successors	Predecessor type	Successor type	Minimum absolute cluster size *(n employees)*	Minimum relative cluster size	
					share in predecessor employment	*share in successor employment*
Type of linkage						
1. Largely identical	1 to 1	any	any	5	50%	50%
2. Absorption	1 to 1	exit	continuing	5	75%	-
3. Split-off	1 to 1	continuing	entrant	5	-	75%
4. Absorption (bis)	1 to 1	exit	continuing	10	50%	-
5. Split-off (bis)	1 to 1	continuing	entrant	10	-	50%
6. Merger of exits	*n* to 1	all exits	entrant	5	50% (1)	50%
7. Break-up into entrants	1 to *n*	exit	all entrants	5	50%	50% (2)
8. Merger other	*n* to 1	any	entrant	5	-	25% (3), 50% (4)
9. Break-up other	1 to *n*	exit	any	5	25% (3), 50% (4)	-
10. Cluster >= 30	1 to 1	any	any	30	10%	10%

(1) Share of the sum of the clusters in total employment of the predecessors
(2) Share of each individual cluster in employment of successor
(3) Share of each individual cluster
(4) Share of the sum of the clusters

The types of linkages are not mutually exclusive. Some of them considerably overlap. Column (1) of Table 2.A.6 shows the total share of linkages that is covered by each type. Column (2) shows the share of additional linkages by type.

Table 2.A. 6 Share of employee-flow linkages by type

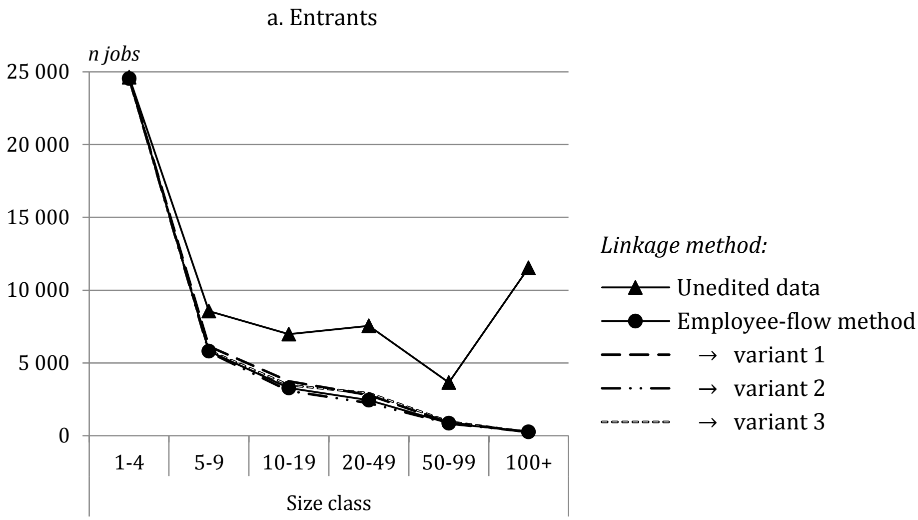
		Separate share in total (1)	Additive share in total (2)
1.	Largely identical	0.57	0.57
2.	Absorption	0.29	0.22
3.	Split-off	0.17	0.11
4.	Absorption (bis)	0.20	0.02
5.	Split-off (bis)	0.11	0.01
6.	Merger of exits	0.05	0.02
7.	Break-up into entrants	0.02	0.01
8.	Merger other	0.04	0.01
9.	Break-up other	0.02	0.00
10.	Cluster >= 30	0.18	0.03
Total		1.00	1.00

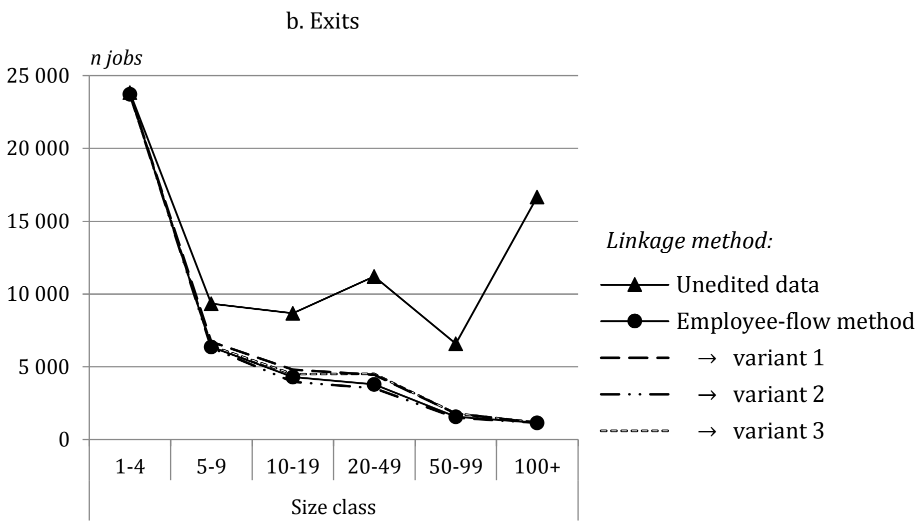
Note: Based on quarterly employee-flow links over the 2003-2012 period.

Robustness checks

The relative size thresholds imposed in the decision rules presented above are to a certain extent arbitrary, and the set of minor rules that define type 4 to 10 may be somewhat confusing. Both are however not critical to the empirical results. Several robustness checks have been performed to test the sensitivity of the results to the set of conditions imposed by the linkage algorithm. Variant 1 is a more restrictive version of the linkage procedure in which all relative cluster size thresholds are increased by 25 percent. Variant 2 relaxes the rules by decreasing the relative thresholds by 25 percent. Variant 3 is again more restrictive and applies only linkage types 1, 2, 3, and 10. The three variants lead to only marginal changes in the empirical estimates compared to the base-line employee-flow results. This is illustrated in Figure 2.A.3, which resumes one of the empirical measures that is most sensitive to missing linkages. The straight lines show the employment distributions at entry and exit, as given by the unedited data and by the base-line employee-flow method (cf. Figure 2.1 in the main text). The dotted lines present the distributions based on the three variants of the employee-flow method.

Figure 2.A. 3 Employment distribution of entrants and exits. Robustness checks of the decision rules of the employee-flow method

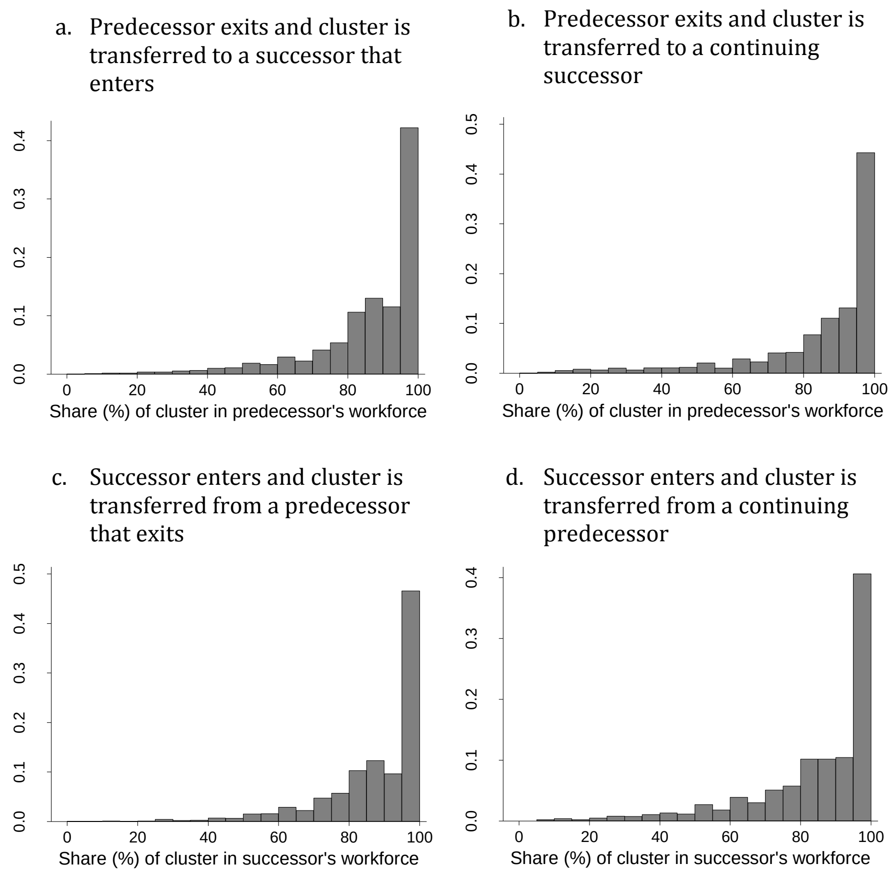




Note: Annual averages over the 2003-2012 period.

Both figures illustrate that increasing the size thresholds by 25 percent (variant 1) removes only few of the original employee-flow linkages and hardly affects the empirical results. The right-skewness of the employment distributions is preserved and total job reallocation by entrants and exits is only slightly increased relative to the base-line employee-flow results (by +3% and +4% respectively). Relaxing the thresholds (variant 2) has an even smaller effect (-1% and -2% respectively), and also variant 3 barely affects the results (+2% and +3% respectively).

Figure 2.A. 4 Frequency distribution of firms by share of employee cluster in firm's workforce



Note: Annual averages over the 2003-2012 period.

The reason why considerable changes in the decision rules hardly affect the results is because the employee clusters that link two firm identification numbers mostly represent close to 100 percent of the workforce of the predecessor, the successor, or both. This finding confirms Benedetto et al. (2007). Firms that change ID number, are absorbed or split-off indeed usually continue operations with mostly the same workforce, apart from naturally in and outflow of individual employees. Figure 2.A.4 illustrates this for the most common types of events. Panel a. presents the distribution of all firm ID numbers that exit the dataset and from which a cluster of at least 5 employees is transferred to an entering ID number. The firms are distributed by the relative size of the clustered employee flow. For 42 percent of the firms, the cluster represents more than 95 percent of the workforce, and for 77 percent of the firms, the cluster exceeds 80 percent of the workforce. Similar results are reported by the examples in the other panels.

2.C Re-estimating measures of entry, exit and growth

Reclassifying entrants and exits

Improved longitudinal linkages are primarily used to identify continuing firms that are misclassified as entrants and exits. They will be labeled as 'spurious' entrants and exits. As is the common practice, they are removed from the entry and exit populations to obtain improved measurements. Spurious entrants are new identification numbers that are linked to a previously active firm, and spurious exits are discontinued identification numbers linked to a subsequently active firm. Formally, an entrant in a given period from t-1 to t is identified as spurious if it is linked to an active predecessor, i.e. a firm with at least one employee in t-1. Similarly, a spurious exit in period t-1 to t is an exiting firm linked to an active successor in t. Other entrants and exits are labeled as 'real'.

Imputing employment of linked firms

Since we are interested in the effect of improving longitudinal linkages not only on aggregate statistics, but also on firm-level estimates, a consistent solution for revising employment growth of linked firm is required. Different strategies can be followed. We propose a simple solution for imputing employment growth at the firm level. Aggregate statistics then follow naturally from the revised firm-level observations.

The following example illustrates the problem. Suppose a link is identified between two firms that merge into one entity with a new administrative identification number. The two firms, previously misclassified as exits, are now identified as continuing. The jobs of these firms are not lost, neither should the jobs that appear at the new entity be treated as job creation. The approach of the U.S. Bureau of Labor Statistics (Pinkston and Spletzer 2002) and of Statistics Canada (Dixon and Rollin 2012) is to collapsing both firms into a consolidated employer. Employment change at the level of this merged entity will now reflect true job creation or destruction. The disadvantage of this approach is that the firm counts will be inconsistent across time and, more importantly, the relation between firm size and firm growth will be biased. Indeed, the size of the consolidated entity is by construction larger than those of the original firms. The alternative approach adopted in this paper it to impute employment growth of each of the two firms by assuming that their individual growth rates equal the one of the consolidated entity. For aggregate measures of job reallocation, this strategy yields the same result as one in which firm-level observations are replaced by a consolidated entity. The advantage of our approach is that we do not intervene in the number of firms and preserve the firm size distribution at the beginning of each period. This allows for consistent estimations of firm-level measures that depend on firm size, and for a direct comparison of firm-level growth rates before and after applying the linkage procedures.

The approach adopted in this paper is based on the following consideration: if firms change identification number, merge or split-up, existing jobs are administratively transferred between ID numbers and it would be wrong to classify employment disappearing at one ID number as job destruction and employment appearing at the other ID number as job creation. At the same time, the firms involved can expand or contract, and thus actually create or destroy jobs. Therefore, job reallocation between firm identification numbers involved in the same event K should not be considered as job creation and destruction, but job growth or decline at the aggregate level of K should be.

Examples of the three most frequent types of events - ID changes, split-offs and take-overs - may serve to illustrate this point. If firm A increases employment from 10 to 12 employees and changes identification number from i to j in the same period, a missing link between the two identification numbers would generate spurious job destruction at the level of the disappearing identification number i and spurious job creation at the new identification number j. It is obvious that firm-level employment growth of A should be calculated as the net employment change at the sum of both identification numbers (+2 employees), and that excess job reallocation at the level of the individual identification numbers should be eliminated. Similarly, a take-over of i by j would generate spurious job destruction at the level of i and spurious job creation at the level of j, which would be recorded as an expanding firm. Firm j, however, increases employment through the acquisition, and not by creating new job opportunities. Just as i should not be regarded as an exit destroying jobs, the jobs reallocated to the acquirer are not to be considered as job creation. Employment growth at the level of the sum of both units, however, does correspond to actual job creation (or destruction). As a final example, consider firm *i* which splits off part of its activities into a new legal entity *j*. The jobs that are split off are not lost, nor does the new entity create new employment. Yet aggregate employment growth or decline at the level of *i+j* should be taken into account.[24]

The revision of employment is implemented by imputing firm-level employment in t, assuming the same growth rate from t-1 to t for each firm involved in the same event. Formally, let K be an event consisting of n interlinked firms i in the period from t-1 to t, then imputed employment $\widetilde{E_{it}}$ of firm i in t equals

$$\widetilde{E_{it}} = E_{it-1} * \frac{\sum_{i=1}^{n} E_t}{\sum_{i=1}^{n} E_{it-1}} \quad (1)$$

where E_i is administratively recorded employment of $i = 1, \dots n \in K$

Firm-level growth is then calculated as the difference between imputed employment in t and registered employment in t-1. It is clear that employment in t at the level of the event K is not affected by the imputation procedure since $\sum_{i=1}^{n} \widetilde{E_{it}} = \sum_{i=1}^{n} E_t$.

In the case of most events, equation (1) simplifies to an expression with a straightforward interpretation, as illustrated below.

1. ID change: one-to-one link between an exit *i* and an entrant *j*.
   If firm *A* changes identification number from *i* to *j* in period t-1 to t, then *i* will be considered a continuing firm for which imputed employment in t is given by *j*:
   $\widetilde{E_{it}} = E_{jt}$ and $\widetilde{E_{jt}} = 0$.

2. Split-off: one-to-one or one-to-many link between a continuing firm *i* and *n* entrants *j*.
   If part of firm *i* is split off into one or more new firms *j*, imputed employment of *i* in *t* is calculated as the sum of employment of all units:
   $\widetilde{E_{it}} = E_{it} + \sum_{j=1}^{n} E_{jt}$ and $\widetilde{E_{jt}} = 0$ for each $j = \{1, \dots, n\}$.

[24] Another example is the relocation of activities between legal entities of the same company. Large companies often register parts of the activities in different firm identification numbers, which are either functionally or geographically split up. Reorganizations of activities between these numbers, usually for legal or accounting convenience, result in administrative transfers of employees which generate spurious creation and destruction at the individual identification numbers. Linking the identification numbers at the event-level corrects for this.

3. Take-over: one-to-one or many-to-one link between *n* exits *i* and a continuing firm *j*.
   If one or more firms *i* are taken over by an established firm *j*, then *i* and *j* are considered continuing firms for which imputed employment in *t* is given by their share in the sum of all units in *t*-1:
   $\widetilde{E_{kt}} = \frac{E_{kt-1}}{\sum_{i=1}^{n} E_{it-1}+E_{jt-1}} * E_{jt}$ for each $k = \{1, \dots, n, j\}$.

4. One-to-one link between two continuing firms *i* and *j*.
   If two continuing firms *i* and *j* are linked, imputed employment of *i* and *j* in *t* is given by the share of each firm in the sum of both units in *t*-1:
   $\widetilde{E_{kt}} = \frac{E_{kt-1}}{E_{it-1}+E_{jt-1}} * (E_{it} + E_{jt})$ for each $k = \{i, j\}$.

5. Merger of exits: many-to-one link between *n* exits *i* and an entrant *j*.
   If several exits *i* are merged into a new firm *j*, each *i* is considered a continuing firm for which imputed employment in *t* is given by its share in the sum of all *i* in *t*-1:
   $\widetilde{E_{it}} = \frac{E_{it-1}}{\sum_{i=2}^{n} E_{it-1}} * E_{jt}$ for each $i = \{2, \dots, n\}$ and $\widetilde{E_{jt}} = 0$.

6. Break-up into entrants: one-to-many link between an exit *i* and *n* entrants *j*.
   If an exit *i* is broken up into several new firms *j*, imputed employment of *i* in *t* is calculated as the sum of employment of all *j* in *t*:
   $\widetilde{E_{it}} = \sum_{j=2}^{n} E_{jt}$ and $\widetilde{E_{jt}} = 0$ for each $j = \{2, \dots, n\}$.

7. Merger of parts of firms: many-to-one link between *n* continuing firms *i* and an entrant *j*.
   If parts of several continuing firms *i* are split off and merged into a new firm *j*, imputed employment for each *i* in *t* is given by its share in the sum of all *i* in *t*-1:
   $\widetilde{E_{it}} = \frac{E_{it-1}}{\sum_{i=2}^{n} E_{it-1}} * \left(\sum_{i=2}^{n} E_{it} + E_{jt}\right)$ for each $i = \{2, \dots, n\}$ and $\widetilde{E_{jt}} = 0$.

8. Break-up and take-over: one-to-many link between an exit i and *n* continuing firms *j*.
   If several parts of an exit *i* are taken over by different continuing firms *j*, then *i* and *j* are considered continuing firms for which imputed employment in *t* is given by their share in the sum of all units in *t*-1:
   $\widetilde{E_{kt}} = \frac{E_{kt-1}}{E_{it-1}+\left(\sum_{j=2}^{n} E_{jt-1}\right)} * \sum_{j=2}^{n} E_{jt}$ for each $k = \{i, 1, \dots, n\}$.

9. In the case of a more complex combination of interlinked firms, the general formula presented in equation (1) applies.

Imputation of employment is performed on a year-by-year basis, i.e. for firms involved in an event in a given period from t-1 to t. In the next period from t to t+1, we restart from registered employment in t and impute employment in t+1 for events in that period. Geurts and Van Biesebroeck (2014) have extended the imputation method over a five-year period. They found that firms involved in an event are more likely to be involved in another event thereafter. Reconstructing longitudinal firm histories over several years thus rapidly becomes a complex exercise as multiple changes in identification numbers have to be taken into account.

2.D Firm-level growth estimates

To document the relationship between firm size and firm growth, we regress annual net employment growth at the firm level on firm size classes using a saturated dummy regression model that is estimated by OLS. As explanatory variables we include six size classes (1-4, 5-9, 10-19, 20-49, 50-99, and 100 or more employees), as well as industry and year dummies. The model includes a separate indicator for all possible values taken by the discrete explanatory variables, as well as the set of interactions between the size and industry dummies. The dependent variable is calculated as the discrete-time firm-level growth rate using average employment size in year t-1 and t in the denominator, as proposed by Davis et al. (1996a). Denoting employment of firm i in year t as E_{it}, its growth rate over the preceding year is $g_{it} = (E_{it} - E_{it-1})/\bar{E}_{it}$, 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.[25]

We follow Davis et al. (1996b) for the size classification of continuing firms, which is based on average employment in year t-1 and t instead of on base-year employment in year t-1.[26] This classification is used to mitigate regression-to-the-mean effects from a traditional base-year approach. Exits are assigned to the size class of employment in their last year. Finally, as in Haltiwanger et al. (2013), we

[25] Some continuing firms have zero employment in t-1 or t ('dormant' firms). They are treated as outliers and are omitted from the estimations in the period concerned.

[26] Due to averaging and employment imputation for transfers, employment is a continuous variable. Firms are classified into the following size intervals:]0,5[, [5,10[, [10,20[, [20,50[, [50,100[, [100,$\propto$[.

estimate employment-weighted specifications of the model, which enables the size coefficients to be interpreted as net employment growth rates for a given size class of firms.

The following regression model is estimated:

$$g_{it} = \sum_{k=1}^{6}\left(\alpha_k + \beta_k^d\, D_{it}^d]\right)1[size_{it} = k] + \sum_d \gamma_d\, D_{it}^d + y_t + \varepsilon_{it} \qquad (2)$$

where g_{it} is the growth rate of firm *i* in the period from t-1 to t and the dummy variable $1[size_{it} = k]$ takes a value of one if the average size of firm *i* in period t-1 to t equals k. The industry dummies D_{it}^d enter both additively and interacted with the set of size dummies. As we enforce $\sum_d \beta_k^d = 0$, the average effect of size on growth is captured by the uninteracted α_k coefficients, while the β_k^d coefficients allow for industry heterogeneity. The additive year dummies y_t control for business cycle effects.

2.E Job reallocation measures

To document aggregate measures of job creation and destruction, we follow Davis et al. (1996a), and decompose the net employment growth rate into the job creation rates by entry and by expansion, and the job destruction rates by exit and by contraction. Formally, if E_t is total employment in year t, and g_t is the net employment growth rate between year t-1 and t, then

$$g_t = \frac{(E_t - E_{t-1})}{X_t} = \frac{JR_t^{entry}}{X_t} + \frac{JR_t^{expand}}{X_t} - \frac{JR_t^{exit}}{X_t} - \frac{JR_t^{contract}}{X_t} \qquad (3)$$

where $X_t = (E_t - E_{t-1})/2$ is the average employment in year t-1 and t

and JR_t^{entry}, JR_t^{expand},JR_t^{exit},$JR_t^{contract}$ are the number of jobs created by entrants and expanding firms, and destroyed by exits and contracting firms respectively in period t-1 to t. Each component $\frac{JR_t^i}{X_t}$ represents the job creation or destruction rate of the corresponding group of firms.

2.F Job reallocation by industry

The general results discussed for the total private sector are also found at the disaggregated industry level, but the biases are exacerbated in industries with a relatively high share of large firms. This is most noticeable in manufacturing and in the sector of human health and social work, where average firm size is much larger than in other industries (29.4 and 33.5 employees respectively). Table 2.A.7 shows that the traditional method produces job reallocation rates by entry and exit, which are overestimated by up to 200 percent compared with benchmark results. The employee-flow method, which more easily identifies missing linkages in larger size classes, reduces the biases to less than 10 percent.

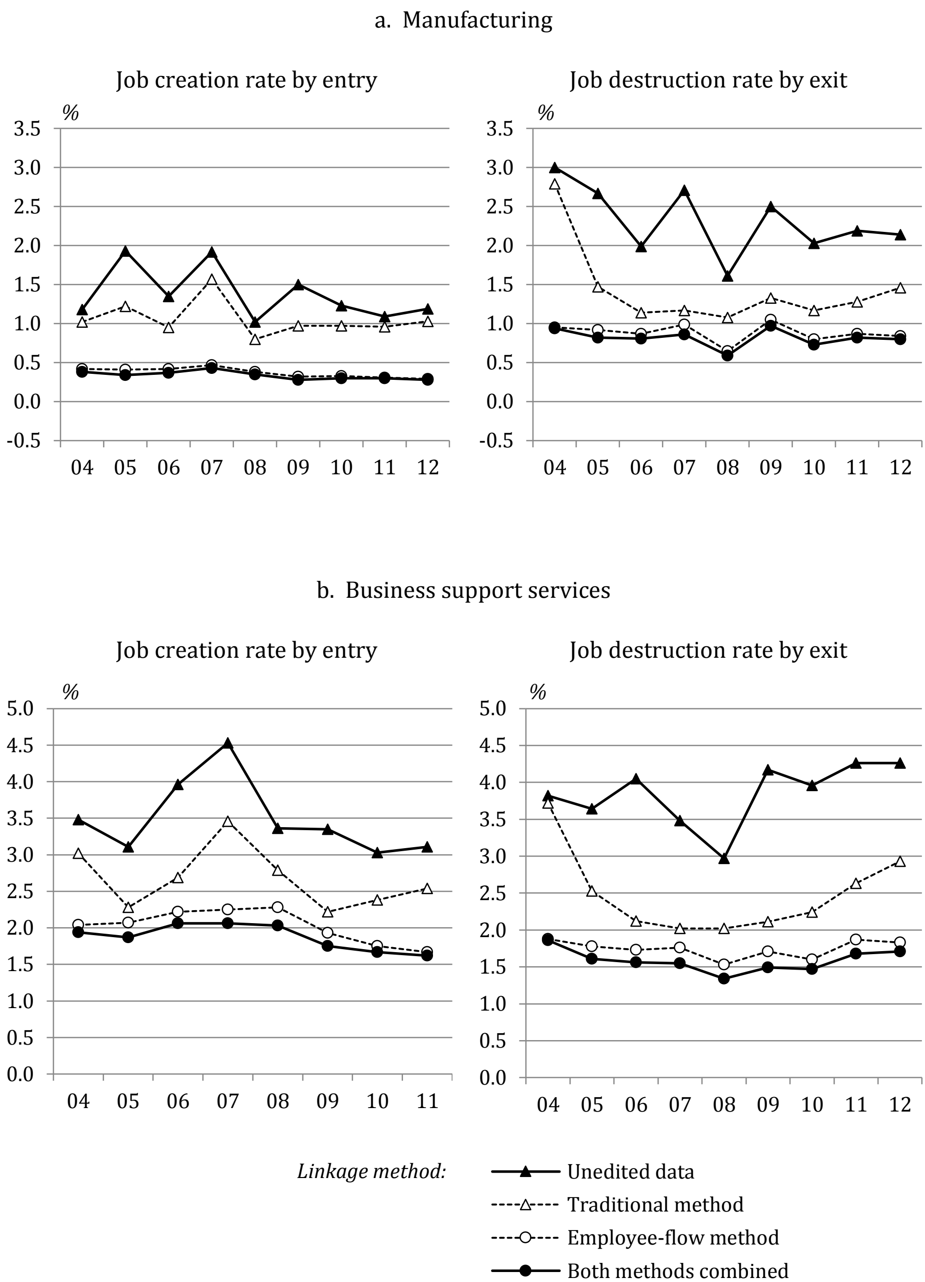
In line with the patterns of annual variation discussed in Section 2.5.3, job reallocation rates by entry and exit in most industries are much more stable over time than it is suggested by unedited or partly edited longitudinal data. Figure 2.A.5 shows the results for some selected industries. The unedited data and the traditional method report large fluctuations in the job creation rate by entry, which are often reflected symmetrically in the job destruction rate by exit. The employee-flow method strongly reduces this spurious variation and reveals that in most industries, job reallocation rates by entry and exit vary smoothly in time.

Table 2.A. 7 Annual job reallocation rates by industry

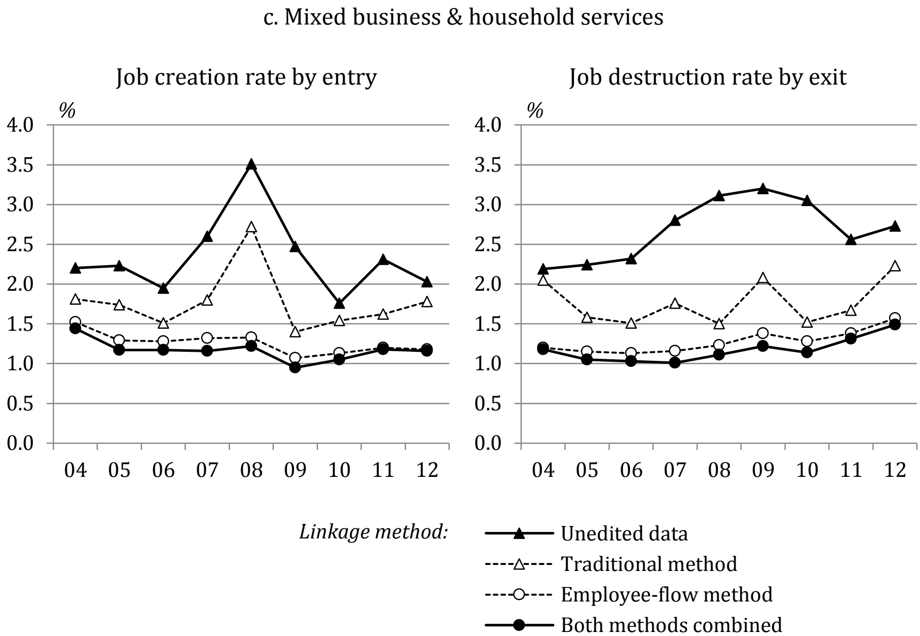
	Rates (%)				*Percent bias vs. both methods combined*			
	Job creation rate		Job destruction rate		Job creation rate		Job destruction rate	
	By entry	By expan-sion	By exit	By contrac-tion	By entry	By expan-sion	By exit	By contrac-tion
Industry (net growth rates in parentheses)								
Agriculture (+0.2%) - *average firm size: 5.1*								
Unedited data	6.1	15.7	7.4	14.3	*25*	*0*	*20*	*1*
Traditional method	5.3	15.9	6.5	14.1	*8*	*1*	*5*	*-1*
Employee-flow method	5.3	15.6	6.6	14.2	*9*	*0*	*8*	*0*
Both methods combined	4.9	15.7	6.2	14.2				
Manufacturing (-1.3%) - *average firm size: 29.4*								
Unedited data	1.4	4.5	2.3	4.9	*309*	*29*	*184*	*13*
Traditional method	1.1	3.9	1.3	4.8	*214*	*10*	*55*	*11*
Employee-flow method	0.4	3.6	0.9	4.4	*10*	*2*	*8*	*1*
Both methods combined	0.3	3.5	0.8	4.3				
Construction (+1.3%) - *average firm size: 7.9*								
Unedited data	3.4	8.6	4.4	6.3	*40*	*9*	*51*	*3*
Traditional method	2.9	8.3	3.5	6.2	*21*	*5*	*22*	*1*
Employee-flow method	2.7	8.0	3.2	6.2	*9*	*1*	*9*	*0*
Both methods combined	2.4	7.9	2.9	6.1				
Trade (+0.8%) - *average firm size: 8.7*								
Unedited data	2.6	6.5	3.2	5.0	*67*	*12*	*78*	*7*
Traditional method	2.0	6.0	2.3	5.0	*30*	*4*	*27*	*6*
Employee-flow method	1.7	5.9	1.9	4.8	*7*	*1*	*9*	*1*
Both methods combined	1.6	5.8	1.8	4.7				
Accommodation and food services (+0.8%) - *average firm size: 5.8*								
Unedited data	7.0	11.6	6.8	10.9	*28*	*5*	*36*	*3*
Traditional method	6.5	11.1	6.2	10.8	*19*	*0*	*24*	*2*
Employee-flow method	5.6	11.1	5.2	10.7	*3*	*0*	*3*	*0*
Both methods combined	5.5	11.0	5.0	10.6				
Business support services (+3.5%) - *average firm size: 11.4*								
Unedited data	3.4	9.2	3.8	5.3	*87*	*17*	*143*	*15*
Traditional method	2.6	8.6	2.3	5.1	*42*	*9*	*47*	*11*
Employee-flow method	2.0	8.1	1.7	4.8	*7*	*2*	*10*	*4*
Both methods combined	1.8	7.9	1.6	4.6				
Mixed business & household services (-0.3%) - *average firm size: 12.2*								
Unedited data	2.3	4.9	2.7	4.8	*101*	*18*	*129*	*9*
Traditional method	1.8	4.5	1.7	4.8	*51*	*9*	*48*	*10*
Employee-flow method	1.3	4.2	1.3	4.4	*8*	*2*	*9*	*2*
Both methods combined	1.2	4.1	1.2	4.4				
Human health and social work (+3.6%) - *average firm size: 33.5*								
Unedited data	1.4	4.8	1.4	1.3	*171*	*7*	*277*	*18*
Traditional method	1.3	4.6	1.1	1.3	*154*	*2*	*214*	*21*
Employee-flow method	0.5	4.5	0.4	1.1	*7*	*1*	*7*	*5*
Both methods combined	0.5	4.5	0.4	1.1				

Note: Annual averages over the 2003-2012 period. Net employment growth rates by industry are decomposed in a similar way as for the total private sector (see Appendix 2.E).

Figure 2.A. 5 Annual job reallocation by entry and exit in selected industries



(figure 2.A.5 continued)



Note: The job creation (job destruction) rate represents the employment share of entrants (exits) in total employment.

Chapter 3

Firm creation and post-entry dynamics of *de novo* entrants

Abstract

We show that within the same age cohort, growth rates of young firms are strongly increasing in firm size. This robust empirical pattern is confined to the initial years after entry; in line with many previous studies, we find that growth rates become independent of size as a cohort matures. Both the initial pattern and the subsequent convergence are consistent with the framework of the passive learning model if young firms adjust their size only slowly to new information, for example due to financing or hiring frictions. Importantly, we focus our analysis on firms that enter *de novo*. They are defined as new firms starting new operations and hiring their first employee. We distinguish them from pre-existing companies that merely re-register as a new firm, for example following a restructuring or merger. The extremely narrow size distribution that we observe for *de novo* entrants provides further support for the passive learning model.

JEL Codes: L11, L25, M13

Keywords: Firm dynamics; Passive learning model; Growth

This chapter is joint work with Johannes Van Biesebroeck. At the time of publication of this manuscript, the paper is in a review and resubmit mode in the *International Journal of Industrial Organization*. We would like to thank several people for valuable comments and suggestions for improvement: Eric Bartelsman, John Earle, Joep Konings, Otto Toivanen, three anonymous referees and seminar participants at the OECD Workshop Understanding Productivity growth, the Belgian Federal Planning Bureau, the Bank of Finland, the Bank of Italy, Tilburg University, Copenhagen Business School, the Düsseldorf Institute for Competition Economics, CREPP Université de Liège, 2015 ECORE Summer School on Market Power, and the IWH Workshop on Entrepreneurship and the Labour Market. Financial support from ERC grant No. 241127 and KU Leuven Program Financing is gratefully acknowledged.

3.1 Motivation

New firms entering the economy are generally both numerous and small. As an entry cohort matures, the average firm size increases and the size distribution, being initially highly right-skewed, shifts to the right. Empirical studies have consistently documented how selection leads to a rapid increase in concentration in a given cohort: many young firms fail shortly after entry and firms that expand have a higher probability of survival than firms that stay small (Evans 1987a; Dunne, Roberts and Samuelson 1989; Mata, Portugal and Guimaraes 1995). The passive learning model of Jovanovic (1982) has been widely used to rationalize these post-entry patterns. It assumes that firms enter with an innate productivity they do not know themselves at entry but gradually discover by operating in the market. Firms that learn they are more efficient grow and survive, while the inefficient exit.

Less consensus exists how growth patterns of young surviving firms contribute to the tendency towards increased concentration in a given cohort. Empirical studies typically find that growth rates are very high in the first years after entry and rapidly decrease with age, another regularity in line with model of Jovanovic (Evans 1987a; Haltiwanger, Jarmin and Miranda 2013; Mata and Portugal 2004). But it is unclear whether within a cohort smaller firms grow faster and to some extent catch up in size, or whether larger firms have higher growth rates. The first pattern would imply a negative size-growth relationship and slow down concentration, while the second pattern would accelerate the trend towards increased concentration (Dunne and Hughes 1994).

Knowing the form of this relationship is important for two reasons. First, theoretical models of firm dynamics often assume or imply a specific relation between growth and size. Second, policy measures to support entrepreneurship and growth often discriminate between firms of different size.

The few studies that have examined the relationship between growth and size of young survivors conditional on age, both measured in terms of employment, report contrasting findings. Evans (1987a), Lotti, Santarelli and Vivarelli (2003) and Mata (1994) find a negative relationship, but Haltiwanger et al. (2013) conclude that there is no systematic relationship between firm size and growth. Furthermore, when using their preferred methodology, Haltiwanger et al. (2013) find that the growth-size relationship within a given age cohort is positive, both for young and older firms.[1] The model of Jovanovic provides little guidance either. In the general version, the relationship is undetermined. Only under specific assumptions does the model predict growth among firms of the same age to be independent of size.

We use data for the universe of Belgian employer firms over a ten-year period and find that the size-growth relationship of young, surviving firms of the same age is strongly and robustly positive. We show, however, that this relationship is confined to the very first years of operation. When entrants mature, the empirical pattern converges to growth rates that are more or less proportionate to size. This convergence confirms previous studies showing that growth rates are independent of size for older and larger firms (Mansfield 1962; Hall 1987; Geroski 1995). A positive size-growth pattern among older firms, as in Haltiwanger et al. (2013), cannot be a steady state as the firm size distribution would become degenerate.

Both the initial deviation and the subsequent convergence are consistent with an augmented passive learning model. The initial pattern can be rationalized within the Jovanovic (1982) framework if one takes into account that young firms adjust their size only gradually to new information and not instantaneously as is assumed in the stylized setting of the model. For example, financing or hiring constraints may prevent young firms from expanding immediately to their desired size. Recent evidence indeed suggests that young firms face more severe financial constraints than older firms (Cabral and Mata 2003; Beck et al. 2006), and that fast-growing firms experience the greatest constraints to growth (Brown, Earle and Morgulis 2015). A similar delay before weak performers exit the industry will also reduce the growth rate of small firms and contribute to the observed positive relationship between growth and size.[2] As firms mature and gradually learn their true efficiency, additional information becomes less informative and they converge to their steady state size.

Two measurement problems we explicitly address are worth highlighting as they illustrate the empirical pitfalls in estimating the relationship between growth and size for young firms. First, the estimated pattern is sensitive to identification of new firms and their post-entry histories in the data. We define *de novo* entrants

[1] The general results presented in Haltiwanger et al. (2013) are shown in greater detail for young firms in Decker, et al. (2014), they confirm the positive relationship.

[2] Abbring and Campbell (2005) show that many poorly performing firms stick around while making losses as they are committed to a year's lease on their premises.

as new firms starting new operations and identify their point of entry as the year they hire their first employee. Second, the potential negative bias in the relationship, induced by sample selection and regression-to-the-mean, is exacerbated in a sample that consists of very small firms. We now discuss these two aspects in some more detail.

Firm-level administrative data are currently the main source for empirical analysis on firm dynamics.[3] The identification of individual firm histories in the data is, however, hampered by the fact that firms may change administrative ID code or restructure, leading to missing links in firm histories over time. It is widely recognized that such events lead to spurious measurements of entry and exit, and to overestimations of firm and employment dynamics (Haltiwanger et al. 2013; Geurts 2016). The bias they introduce in post-entry growth patterns has, however, received less attention. We make use of two state of the art record linking methods to minimize these problems. They enable us to more accurately trace the complete histories of *de novo* entrants, from the moment they hire their first employee till they cease activities, i.e. true economic exit. We distinguish them from spurious entrants, i.e. firms that continue existing economic activities with a new ID code, for example after a merger or split of legal entities, or a change in legal identifier. Similarly, true economic exit is distinguished from firms that disappear from the data without closing down operations. We show that failing to identify even a small amount of spurious entrants has major implications for the estimation of post-entry growth patterns.

Our exclusive focus on *de novo* entrants reveals that the firm size distribution at entry is confined to a much narrower range of small size classes than found in many previous studies. This empirical observation is very much in line with the passive learning model which predicts that firms, lacking prior information about their efficiency, all enter at the same size. Studies that cover a sample with a broad range of firm sizes already at entry must be investigating growth in a different population than *de novo* entrants.

It is well-known that two statistical problems may bias the relationship between size and growth for surviving firms. Regression-to-the-mean as well as sample selection may spuriously induce a negative relationship if firm size is measured in the base year, i.e. at the start of the period over which growth rates are calculated (Hall 1987). Although these problems may be less important for larger firms, the statistical side-effects of the base-year size classification are

[3] See for example cross-country analysis in Bartelsman, Haltiwanger and Scarpetta (2009).

greatly exacerbated in a sample of small firms, as is the case in our sample of *de novo* entrants. We therefore need to directly address these measurement issues. To avoid bias in the size-growth relationship, we use three alternative firm-size classifications that approximate a continuous size-growth relationship. We find a robust positive size-growth relationship for each of the alternatives.

The remainder of the paper is organized as follows. Section 3.2 starts with a brief overview of stylized facts on entry and post-entry dynamics. It also reviews predictions of the model of Jovanovic (1982) and discusses how they are affected by delayed adjustment. Section 3.3 presents the dataset and our strategy to identify *de novo* entrants and their post-entry histories. In Section 3.4, the empirical model and the size measurement issues are discussed. The results are presented in Section 3.5, first showing that some well-established facts about firm entry and exit are replicated in the Belgian dataset, and then showing post-entry growth patterns by age and size. Section 3.6 concludes.

3.2 Facts and theory

3.2.1 Some stylized facts

Empirical studies for various countries have found entry rates of new firms in manufacturing and services to vary between 5 and 15 percent per year. Most entrants tend to be much smaller than the average incumbent, such that the employment share of new entrants is generally far less than 5 percent of the workforce (Siegfried and Evans 1994; Geroski 1995; Caves 1998). As a cohort matures, average firm size increases and the number of firms falls. This tendency towards increased concentration in a given age cohort is very strong in the first years after entry. A typical pattern is that 5 to 10 years after entry, average firm size has doubled, but only half of an entry cohort survives.[4] Cabral and Mata (2003) showed firm size to be highly right-skewed at entry and shift towards a more symmetric distribution over time. The long-run cohort's size distribution remains, however, right-skewed, without convergence to a common size (Konings 1995).

The rapid increase in concentration among an entry cohort is explained by specific post-entry dynamics showing systematic differences between young

[4] See for example Dunne, Roberts and Samuelson (1988) for the U.S., Wagner (1994) and Boeri and Cramer (1992) for Germany; Mata et al. (1995) for Portugal.

firms and incumbents. A first difference is a selection process that reduces the number of smaller firms in a cohort. Many empirical studies have shown that young firms exhibit high failure rates immediately after entry. Two patterns are highly robust: (i) exit rates are decreasing in firm size and (ii) survival rates increase as firms mature.[5]

Another well-established fact is that young surviving firms exhibit remarkably high growth rates which decline with age.[6] Variation in growth rates among surviving firms can contribute to increased concentration if larger (young) firms tend to grow faster than smaller ones, or if growth rates exhibit positive serial correlation (Dunne and Hughes 1994). The existing evidence on which pattern prevails in the early post-entry process has been inconclusive.

Several studies lump all firms below a certain age in one cohort and verify whether growth rates conditional on survival increase or decrease with firm size among these young firms. Dunne et al. (1989) and Almus and Nerlinger (2000) find, for the manufacturing sectors of the U.S. and Germany, that smaller plants or firms grow faster than larger ones. Wagner (1994) also studies German manufacturing firms, but finds growth rates to be independent of size.[7] As these patterns include an age effect within the broader cohort—and we know that younger firms tend to be smaller and growing faster—they provide imperfect evidence on the size-growth relationship among firms of the same age.[8]

The few studies that have investigated post-entry growth conditional on age obtain contrasting results. Evans (1987a) and Lotti et al. (2003) report an inverse relationship between growth and size given age for surviving young firms in the first six years after entry. They find this pattern diminishes with age and converges towards growth that is proportionate with size, consistent with evidence that suggests Gibrat's law holds in a sample of older firms or among firms that have exhausted scale economies (Mansfield 1962; Hall 1987; Geroski 1995). Mata (1994) finds a weak negative relationship that is insignificant at the

[5] See for example Evans (1987a) and Dunne et al. (1989) for U.S. manufacturing plants, Haltiwanger et al. (2012) for U.S. manufacturing and services; Mata et al. (1995) for Portugal.

[6] See the same studies for the U.S.; Mata and Portugal (2004) for Portugal.

[7] The three studies group together all firms younger than, respectively, 5, 6, or 10 years. Audretsch, Santarelli and Vivarelli (1999) also investigate the relation between growth and size of young Italian manufacturing firms, but they estimate growth rates relative to size at entry.

[8] A discussion of this age composition effect is provided in Section 3.5.3.

1 percent level.[9] Haltiwanger et al. (2013) report a negative as well as a positive pattern, depending on the size classification method. When using their preferred methodology, they find larger firms to grow more rapidly than smaller ones among young survivors of the same age. Moreover, their results show no convergence towards proportionate growth for older firms. The contrasting results may be partly explained by differences in measurement methods and industry scope, as we discuss in more detail below.[10] We will show that an accurate identification *de novo* entrants matters greatly too.

Note that some studies have used firms as their unit of analysis while others used plants or establishments. In our analysis, we are not interested in country comparisons of performance, but rather try to uncover general patterns of firm behavior. The unit of analysis most closely related to the theoretical notion of new firm creation is the firm and that is the unit of observation we will work with. As the vast amount of new entrants tend to have only a single plant or establishment, this definition covers a subset of the entrants that plant-level studies would identify.

3.2.2 Theoretical framework

How do firms enter?

The passive learning model of Jovanovic (1982) implies a particular process of firm dynamics by age and size and has often been used to rationalize exit and growth patterns of entrants. The key assumption is that firms enter without knowing their own innate productivity. Prior to entry, they receive, but do not observe, a random draw from the productivity distribution in the industry. Since entrants know the population distribution, they have the same prior beliefs and all enter at the same size.[11] Each period they update their prior distribution over

---

[9] Pooling young firms up to age 4 into one age class, Mata (1994) finds a stronger negative relationship. As noted before, this result is likely to reflect an age composition effect of small, fast-growing firms being younger.

[10] Evans (1987a) and Haltiwanger et al. (2013) report results for U.S. firms. Lotti et al. (2003) cover Italian firms and Mata (1994) Portuguese firms. Haltiwanger et al. (2013) classify firms by average size in t-1 and t and include both manufacturing and services, while the other studies use a base-year size classification and are limited to manufacturing firms.

[11] Models of entrepreneurial entry with financing constraints, such as Evans and Jovanovic (1989) and Cabral and Mata (2003), also predict that the size distribution of entrants will cover a narrow range.

their own productivity level using Bayes' law as evidence on profitability is realized. Firm sizes diverge as the cohort matures even though the underlying firm-specific productivity level is constant.

This modeling approach contrasts with Lucas (1978) which features a dispersion of managerial skill in the population. High-skill individuals self-select into entrepreneurship, rather than becoming an employee, and they choose their firm size optimally upon entry. It also contrasts with the model of Hopenhayn (1992) where firms similarly receive a random draw from a known productivity distribution, but they observe this realization after paying a fixed entry cost and before hiring any production factors. If they enter, they immediately do so at the "right" size.

The first implication of the Jovanovic model rarely holds in large-scale datasets used to investigate firm dynamics. Firms are predicted to all enter at the same or similar scale, while actual entrants typically span a broad range of firm sizes. Deviations might simply be due to the stylized assumptions of the model, but two measurement issues help explain the discrepancy between the prediction and stylized facts. First, variation in entry size can reflect that some time has passed between the moment a firm is established and the first time it is observed in the dataset. In our administrative database of Belgian employers, new firms are observed in the first year they record positive employment on June 30. On that day, some firms have already been in existence, either without employees for an unknown period, or with employees for up to 12 months. They have had the chance to learn about their innate productivity, and choose different growth rates, or even exit, in response. The observed entry size distribution should thus (at least partly) be regarded as the outcome of an initial selection and size adjustment process. For this reason, we denote the first year of entry in the dataset as age 1, and the unknown moment of the firm's establishment as age 0.

Second, and more importantly, the group of entrants in administrative datasets typically includes some pre-existing firms that re-enter the dataset after a legal or ownership restructuring or enter with a new subsidiary. Examples include divestitures, control changes, legal restructuring for tax or liability reason, incumbents entering a new industry or starting activities in a new region, etc. These other modes of entry are certainly economically relevant, but we do not expect post-entry dynamics of these firms to conform to the predictions of the passive learning model. We label them as spurious entrants, as opposed to *de novo* entrants which we study in this paper.[12] Several studies have demonstrated that entry by established firms fundamentally differs from *de novo* entry (Dunne et al. 1988; Baldwin and Gorecki 1987; Konings et al. 1996; Bilsen and Konings 1998; Mata and Portugal 2004). These firms already have a better idea of their own productivity. They tend to enter with a larger size, are less likely to fail, and exhibit less dynamic growth patterns. They are an interesting group of firms to study, as these changes could very well be systematically related to past or future performance, but here we choose to focus on *de novo* entrants.

Just as the optimal size with which firms enter in Lucas (1978) or Hopenhayn (1992), the uniform entry size in Jovanovic (1982) is an extreme assumption in a stylized model of entry. Even *de novo* entrants may possess some pre-entry knowledge about their resources and capabilities which affects both entry decisions and subsequent success (Helfat and Lieberman 2002). Moreover, some entrepreneurs enter small simply because they have no or limited growth intentions (Acs, Astebro, Audretsch and Robinson 2016). Initial size may thus partially reflect both prior knowledge and post-entry growth paths. The substantial size dispersion at entry that is generally observed in empirical datasets has been used to explain post-entry growth patterns (Audretsch et al. 1999). After carefully identifying *de novo* entrants, however, our sample reveals very little variation in entry sizes, reflecting more closely the stylized assumption in Jovanovic (1982).

How do firms grow after entry?

Many heterogeneous firm models do not incorporate firm-specific stochastic elements that give rise to systematic heterogeneity in growth rates. In the model of Hopenhayn (1992), firms enter immediately at their optimal size and later adjustments in firm size are responses to random productivity shocks firms have no control over. Abbring and Campbell (2004) add persistence in post-entry shocks to the model which leads to serial correlation in growth rates and eventually to a positive size-growth relationship.

The passive learning model of Jovanovic (1982) is one exception.[13] Firms only discover their own innate efficiency level from operating in the market. Initially,

---

[12] As shown in Section 3.3., the vast majority of spurious entrants we distinguish from *de novo* entrants are simply incumbents that continue the same activities with a new identification code after an administrative or legal change.

[13] The active learning model of Ericson and Pakes (1995) is another exception. In their model, growth is a function of firms' actions as they can make investments to raise

they have the same beliefs about this and they all enter at the same size. Realized profits depend on their actual underlying efficiency and idiosyncratic cost shocks and they use Bayes' rule to update their beliefs and expand or contract into their correct size. Firms that discover they are more efficient, grow and survive, while the inefficient shrink and exit. As time passes, firm sizes within an entering cohort diverge and become strictly increasing in firms' estimate of their own efficiency. As firms mature and gradually learn their true efficiency, additional information becomes less informative and they converge to their steady state size.[14]

This model generates several testable predictions about exit and growth patterns in relation to the firm's age and size. First, the noisy selection process implies an inverse relationship between exit and size given age and between exit and age. Unsuccessful firms stay small, they might even contract, and eventually choose to exit. Larger firms are those that received favorable cost information in previous periods and have expanded. While initial profit realizations provide new entrants with a lot of information on their ability, subsequent information becomes gradually less informative and is less likely to induce exit.

Second, the model implies that conditional on survival younger firms have higher and more variable growth rates than older firms. They are still highly uncertain about their own quality and respond to market success by expanding. As the weakest firms exit, average efficiency among surviving firms improves from period to period which is reflected in higher average firm sizes. As firms mature, revisions of estimated efficiency become smaller. Firms eventually approach their optimal scale and the variance of growth rates converges to zero.

Third, because smaller firms are on average younger, the model also predicts an inverse relationship between growth rates and size in a cross-section of firms that encompasses a range of cohorts. Several empirical studies find evidence for this inverse relationship and Jovanovic (1982) cites it as a key motivation for the model. However, without additional assumptions, the model does not imply any

---

productivity. As the link between investment and productivity is stochastic, even identical investments can generate different outcomes.

[14] Further growth is driven solely by business cycle shocks affecting all firms similarly. In the 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.

systematic relationship between growth rates and size conditional on age.[15] Assuming a Cobb-Douglas cost function leads to a prediction that growth rates are independent of firm size among firms of the same age cohort, consistent with Gibrat's law. At each point, a firm's size reflects its best estimate of its efficiency. With this cost assumption, adjustment is complete and subsequent adjustments depend only on future information which is by definition random.

In the stylized framework of the Jovanovic model, a firm's current size only reflects its past growth history. The model assumes instantaneous adjustment to new information, but in reality, frictions might distort this process. Hsieh and Klenow (2009) show for several countries that deviations between factor prices and marginal productivities and between observed and optimal output levels are widespread. As these deviations partially reflect the dynamic adjustment of quasi-fixed production factors (Asker, Collard-Wexler and De Loecker 2014), it is likely that younger, less established firms face greater external frictions. For the prediction of the size-growth relationship conditional on age, it matters greatly whether they already affect firm size at the moment of entry or whether they mainly influence adjustments in firm size following entry.

A prominent example of the first situation is the model of Evans and Jovanovic (1989) where entrepreneurs face liquidity constraints. Heterogeneity among firm size at startup reflects that the smallest entrants faced the strongest financial constraint. They need to finance their expansion from realized profits. If the friction is not perfectly correlated with ability, they will also have the highest growth potential and we should observe a negative relation between initial size and subsequent growth, as in Audretsch, Santarelli and Vivarelli (1999). An alternative mechanism that generates the same prediction is developed by Cabral (1995). If production capacity requires substantial sunk costs that are foregone when firms exit, smaller firms are more likely to exit and they will choose to invest gradually and enter at even smaller scale.

In the second situation, entry size is not distorted by frictions. Yet following entry, some firms cannot immediately adjust to their desired size when they revise their estimate of their innate efficiency. Credit, hiring, or regulatory constraints can limit growth in the first years. For some expanding firms, current size will be below desired size and they will need several years to incorporate positive

[15] Dunne et al. (1989) argues that efficiency levels, and thus firm sizes, are bounded from above. This leads to a negative relationship as there is less room for further increases for larger firms.

information into their size. For some years, their size and growth rate both reflect underlying firm quality. Until adjustment is complete and desired size catches up with actual size, it leads to higher growth rates for larger firms. Delayed adjustment of firm size introduces a positive correlation between past and current growth, and thus between firm size and growth.[16]

Delayed adjustment can have many reasons. It can be externally imposed, for example credit constrained firms may need to finance investments from retained earnings. A vast literature documents the excessive sensitivity of many firms' investments to free cash flow (Fazzari, Hubbard and Petersen 1988; Evans and Leighton 1989). Cabral and Mata (2003) and Beck et al. (2006) find that young firms face more severe financial constraints than older firms, while Brown, Earle and Morgulis (2015) show that fast-growing firms experience the greatest constraints to growth. Search frictions to hire specialized staff in thin labor markets or zoning regulations are other external frictions that can delay adjustment to positive shocks. Risk aversion will exacerbate the pattern of gradual adjustment. While larger firms might be risk-neutral, individual entrepreneurs are likely to be somewhat risk averse (Brockhaus 1980). Especially in the face of irreversible investments and sunk costs, firms will not incorporate all positive information immediately in their size. Past growth will result in a somewhat higher size, but also be followed by future growth.

Delayed exit further contributes to a positive relationship between growth and size. The option value associated with the sunk entry costs may provide an incentive for some loss-making firms to continue operations before eventually deciding to withdraw from the market. In many administrative firm-level datasets it is common to observe firms with no employment and no or minimal sales for several years. If fixed costs are low relative to sunk entry costs, small firms might simply hang around for the business cycle to improve rather than exit.

[16] In a Markov Perfect equilibrium, the value of current state variables are sufficient statistics for the entire firm history (Ericsson and Pakes 1995). With adjustment frictions this is not necessarily the case anymore.

3.3 Data

The analysis is based on the register of Belgian employers maintained by the National Social Security Office (NSSO). It includes all private firms with at least one employee and covers the period from 2003 to 2012. In an average year, the sample includes 178 000 firms and 2 070 000 employees.

De novo entrants are defined as new firms starting new operations. We identify their point of entry in the data as the year they hire their first employee. We distinguish them from spurious entrants by making use of two state-of-the-art record linking methods. The methods are further used to trace the complete histories of firms from the moment they start operating till they cease activities, i.e. true economic exit. For those firms that change ID code or restructure, we impute employment measures up to the sixth year of existence. To our knowledge, we are the first to use this approach to obtain consistent post-entry firm histories. The details of our methodological approach are explained in Appendix 3.B. Below, we provide a summary and show that the size range of *de novo* entrants dramatically differs from the size range at entry suggested by the raw dataset. This has major implications for the post-entry size-growth relationship.

It is widely recognized that administrative firm-level data suffer from missing links in individual firm histories, which hinders the straightforward identification of firm dynamics. Firms may change ID code due to mergers, takeovers, split-offs, ownership changes or for tax optimization purposes. These events generate various biases in empirical measures, such as spurious measurements of entry and exit, misclassifications of firm growth across age and size classes, and overestimations of job and firm turnover (Haltiwanger et al. 2013; Geurts 2016). To minimize these problems, we use two record linking methods cumulating the missing linkages we identify.

The first consists of a set of traditional record linking techniques developed by Statistics Belgium in line with the OECD-Eurostat recommendations on constructing longitudinal business data (Eurostat-OECD 2007). The method relies on probability-based matching and the use of supplementary data sources with information on firm continuity. The second linking method is based on an employee-flow approach. It follows one of the key production factors of the firm, the stock of employees, to identify changes in ID codes and firm structure. Continuity of the firm's workforce is thus used to identify firms that operate continuously.

The established linkages are first used to identify continuing firms that are misclassified as exits and entrants in consecutive years. They are labeled as 'spurious' exits and entrants as opposed to true exits and *de novo* entrants. It is especially important to recognize that spurious entrants are pre-existing firms that are likely to exhibit characteristics similar to other incumbents. If they are mixed up with *de novo* entrants, the typical size and growth patterns of young firms will be biased towards those of incumbents. Panel b. of Table 3.A.1 in the Appendix shows that 78 percent of the spurious entrants we identify are simply incumbents that continue the same activities with a new identification code after a purely administrative or legal change. Another 18 percent are split-offs of another firm.

Next, the linkages are used to trace the employment histories of *de novo* entrants that are involved in an ID change or restructuring in the years following entry. When a firm changes ID code its employment history in the data appears to be discontinued. Similarly, firms that merge or split up are recorded with artificial jumps in employment which do not correspond to the actual creation or destruction of jobs. For these firms, we impute employment up to the sixth year after entry. Our approach is to construct an aggregate event-level that includes all firm ID's interlinked in a given period t-1 to t. Firm-level employment in t and t+n is then imputed by assuming the same growth rate for each firm involved in the event. For one-to-one ID changes, which represent the vast majority of events, this simply means replacing the new by the old ID code. For firms that split-up, the method reduces to keeping the entities combined in one firm as before the event. For mergers and more complex events, the firms are kept separated as before the event, and employment of each of them is assumed to exhibit the same grow rate as the merged entity recorded in the data. An important advantage of this imputation method is that it preserves the firm size distribution in t-1 to calculate growth rates from t-1 to t and in subsequent periods, allowing a more accurate estimate of post-entry employment patterns by size.

Table 3.1 shows that the two linkage methods are complementary for the identification of *de novo* entry across different size classes of firms. The first row reports the average annual number of entrants as observed in the raw administrative data. The next rows present the fraction of these firms that are identified as either *de novo* or spurious entrants.

Table 3.1 Share of *de novo* and spurious entrants in all administratively recorded entrants

	Total	By firm size class					
		1-4	5-9	10-19	20-49	50-49	100+
Number of firms	*17283*	*15 368*	*1 209*	*446*	*190*	*39*	*32*
Share of *de novo* entrants	0.91	0.95	0.64	0.41	0.26	0.17	0.03
Share of spurious entrants							
Identified by							
Both methods combined	0.09	0.05	0.36	0.59	0.74	0.83	0.97
Traditional method	0.06	0.05	0.12	0.16	0.21	0.32	0.44
Employee-flow method	0.05	-	0.32	0.57	0.72	0.82	0.97

Note: Average of annual shares over the 2003-2012 period. Firm size classes are based on employment.

Spurious entrants only represent 9 percent of the total, but this low fraction does not mean it is an unimportant group. The probability that a new ID code corresponds to spurious entry increases dramatically with size. They account for more than one third of administrative entrants with 5 to 9 employees and even two thirds of those with 10 or more employees. *De novo* entrants with more than 50 employees are extremely rare. As a result, the size-distribution of *de novo* entrants is more strongly right-skewed than in the unedited data and the presence of spurious entrants would introduce a bias in post-entry patterns by size. Table 3.1 further shows the complementarity of the two linkage methods. The traditional method is needed especially in the size class below five employees, where employee-flow links are absent by construction. Yet the employee-flow method is essential in larger size classes, where it identifies two to three times more spurious entrants than the traditional method.

3.4 Empirical model

We characterize survival and growth patterns for young firms by age and size using the employment history of *de novo* entrants up to the moment of true economic exit. As shown in Dunne et al. (1989), the mean growth rate of a class of firms can be decomposed into the growth rate of survivors weighted by the probability of survival, minus the probability of exit. The two equations, using the firm-level growth rate and the exit dummy as dependent variables, are estimated separately.

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 after June 30 of year t-1 and survive until June 30 of year t. It conditions on surviving a first selection process, 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. Exits in observation period t-1 to t are firms for which t-1 is the last year of positive employment. Firms that change ID code or firm structure are not considered as exits. Their growth path following the event is based on imputed employment. The years between entry and exit, firms are denoted as survivors.[17]

Following Davis, Haltiwanger and Schuh (1996a), firm-level growth rates are calculated as discrete-time employment changes relative to the average of employment in year t-1 and year t. Denoting employment of firm i in year t as E_{it}, the 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.[18] Regressions use employment weights such that the coefficient estimates are readily interpreted as aggregate employment changes for a class of firms. Specifically, the mean estimated growth rate represents the rate of net

[17] Some survivors have zero employment in a given year ('dormant' firms). They are treated as outliers and omitted from the regressions in the periods concerned.

[18] This growth rate is close to the more commonly used logarithmic growth rate $g_{it} = \ln(E_{it}/E_{it-1})$, especially for values between -1 and +1. Both measures show expansion and contraction symmetrically, whereas the growth rate relative to base-year employment t-1 ranges from -1 to infinity. Symmetry is a crucial feature for estimating mean growth rates of young firms, as their employment fluctuates widely. A further advantage of our growth rate is that using the corresponding employment weights, $\bar{E}_{it}$, in the regressions yields coefficient estimates that exactly represent net employment growth of a class of firms. Equivalent weights do not exist for the logarithmic growth rate. In the exit regressions we use E_{it-1} as employment weights.

employment creation in a given age-size class of firms, and the exit rate represents the job destruction rate.

At each age, firms are grouped into six size classes, based on the number of employees and defined on a logarithmic scale:]0,2[, [2,4[, [4,8[, [8,16[, [16,32[, and [32,∝[.[19] All observations with more than 32 employees are in the same size class because few *de novo* entrants reach this size within the first five years of existence. Exits are assigned to the size class of employment in their last year.

To document patterns of firm dynamics, we regress the dependent variables on age and size classes using a saturated dummy regression model. It includes separate indicators for all possible values taken by the two discrete explanatory variables and their interactions. This approach follows Haltiwanger et al. (2013) and has two advantages over other estimation methods used to examine the relationship between growth and size. First, as emphasized by Angrist and Pischke (2009), a saturated regression model fits the conditional expectation function perfectly, regardless of the distribution of the dependent variable. Moreover, no particular shape of the size-growth relationship has to be imposed. Second, the estimates are robust to heteroscedasticity, a recurrent problem in empirical studies of the size-growth relationship.[20]

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

$$y_{it} = \sum_{j=2}^{6} \sum_{k=1}^{6} \left(\alpha_{jk} + \beta_{jk}^{d} \, D_{it}^{d}\right) 1[age_{it} = j] \, 1[size_{it} = k] + \sum\nolimits_{d} \gamma_d \, D_{it}^{d} + \gamma_t + \varepsilon_{it}$$

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 six industry dummies D_{it}^{d} enter both additively and interacted with the full set of age-size interactions. As we impose that $\sum_d \beta_{jk}^{d} = 0$, the average effect of age and size on growth and exit is captured by the uninteracted α_{jk} coefficients, while the β_{jk}^{d}

[19] Due to the use of average employment and imputed employment levels, size is a continuous variable.

[20] For a further discussion of the econometric problems see Hall (1987), Evans (1987b), and Dunne et al. (1989). Since we examine how growth rates of survivors depend on the current size of the firm, where both growth and size are updated at each age, we also avoid the sample censoring bias many previous studies had to address (Mansfield 1962).

coefficients allow for industry heterogeneity. The additive year dummies control for business cycle effects.

Size classification of surviving firms

We approximate a continuous size-growth relationship using three alternative approaches to allocate surviving firms in a size category. The objective is to mitigate two statistical side-effects of a conventional base-year classification, which classifies firms by size in t-1. First, as discussed extensively in the literature, regression-to-the-mean may spuriously induce a negative relationship between size and growth if firm size is measured at the start of the period over which growth rates are calculated. Even if employment growth is independent of size, random variation due to measurement error or transitory fluctuations will systematically bias growth estimates upwards for firms that are small in t-1 (Hall 1987; Friedman 1992; Davis et al. 1996b). Second, employment in the subset of surviving firms is bounded from below by one. Therefore, the lower tail of possible rates of decline is truncated, while the upper tail of growth rates is unaffected. It especially affects smaller firms which will already exit when hit with a moderate negative shock and leads to sample selection bias. It again induces an inverse relation between size and growth if size is determined at the start of the period (Mata 1994; Baldwin and Picot 1995).

Hall (1987) and others have found that these problems have little effect on the size-growth relationship for larger, more established firms. However, they are exacerbated in a population of predominantly small firms, as in our sample of *de novo* entrants. Single employee firms that survive cannot even have a negative growth rate. Dunne et al. (1989) and Mata (1994) largely circumvent these statistical problems by excluding the smallest firms from their sample. This is not an option for us, given our focus on *de novo* entrants which are predominantly observed in size classes below 5 employees.[21] Instead, we use three alternatives to allocate firms in a given size class. The intention is to approximate firm growth in continuous time and we refer to the 'current' size of the firm. A more detailed discussion of these methodologies is in Appendix 3.C; here we provide a brief overview.

The first size classification method, and the one we use for our benchmark estimates, allocates employment gains and losses to each of the size classes that

[21] Among *de novo* entrants, 94 percent of firms have fewer than 5 employees at age 2 and 82 percent at age 6.

the firm passes through as it grows or contracts (Butani et al. 2006). In this 'dynamic' size classification, firms are initially assigned to a size class based on employment in t-1, but are re-assigned to a new class when they cross a threshold. The growth from E_{it-1} to the threshold is assigned to the initial class and the remaining growth from the threshold to E_{it} is assigned to the next size class. This methodology approximates instantaneous class re-assignment that would be feasible if size and growth were measured in continuous time. As it attributes symmetric employment changes to the same size classes, it avoids the negative as well as the positive bias in the size-growth relationship that afflict other methodologies.

The second classification method uses each firm twice in the regression, assigning a weight of one half to each observation. One observation uses the firm's employment level at the beginning of the period—both as a base for the growth rate and to determine the size class. The second observation uses the firm's employment at the end of the period again for both calculations. This approach was proposed by Prais (1958) to avoid regression-to-the-mean bias and can be motivated similarly as the use of average wage shares in a Solow residual, i.e. as a discrete approximation to the continuous Divisia index of productivity growth (Caves, Christensen and Diewert 1982).

A last classification method follows Davis et al. (1996a, 1996b) and uses the average of firm size in years t-1 and t as a proxy for the size over the intervening period. It is adopted for comparison with the results reported by Haltiwanger et al. (2013). Baldwin and Picot (1995), however, indicate that this size classification introduces an upward bias between size and growth if there is positive trend growth rate in the population. [22]

[22] The weights in the growth regressions follow naturally from the three size classification approaches. They always equal the employment used in the denominator of the growth rate calculation: (i) the truncated average employment within the size class, (ii) E_{it-1} or E_{it}, and (iii) $\bar{E}_{it}$.

3.5 Results

In constructing the dataset, we have taken great care to only identify firms as *de novo* entrants when they start new operations, corresponding to firm creation in Jovanovic (1982). With continuing firms misclassified as entrants or exits filtered out, we find two novel patterns. In particular, we show that *de novo* entry is confined to a much narrower range of small size classes than usually found and that growth rates for surviving entrants are increasing with firm size. We discuss the two novel results in detail below, but first summarize a few patterns for *de novo* employer entrants in the Belgian private sector that are consistent with the empirical evidence from other countries, as discussed in Section 3.2. They suggest that the novel findings are not an artifact of the Belgian dataset. A brief summary of the confirmed patterns is provided below, while Appendix 3.D contains a more detailed discussion.

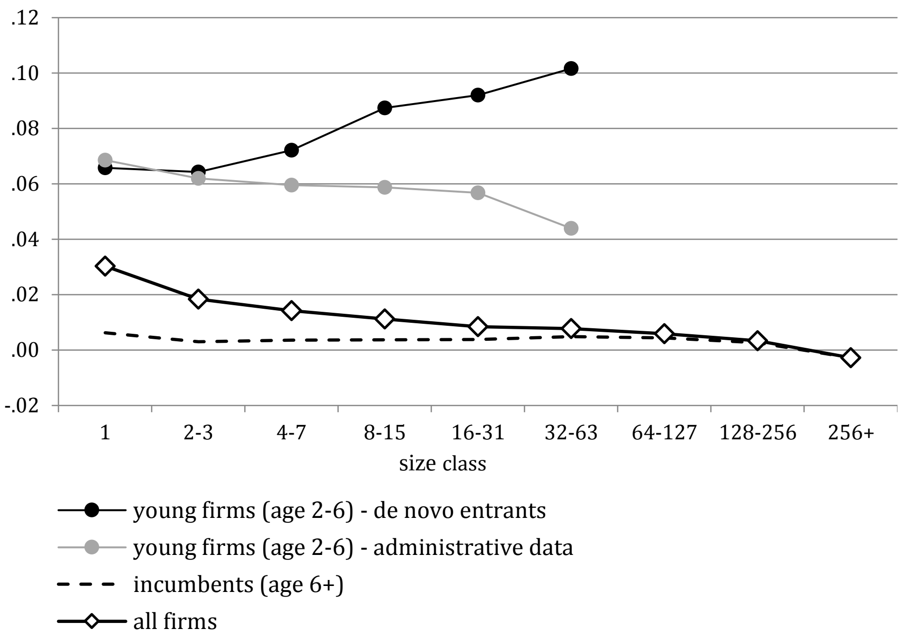
3.5.1 Confirmed patterns

In line with results for many other countries, summary statistics in Table 3.A.2 show that the annual rate of firm entry in Belgium is high (9%), but involves only a small fraction of employment (1.5%). Most entrants are extremely small; average entry size is 1.9 employees, six times smaller than the average size of incumbents. In the years following entry, a large fraction of the entering cohort exits and the average size among survivors increases. Only half of all entrants survive to age 6, at which time the average firm size in the surviving group has almost doubled.

A first mechanism generating this pattern of increased concentration within an entry cohort is selective survival. In line with previous evidence we find high exit rates for young firms which are decreasing in age as well as in size, see panel a. of Figure 3.A.1. Our results suggest that the selection process of the passive learning model—which predicts market exit of the least efficient and therefore the smallest firms—unfolds quickly in the first years after entry. By age 6, exit rates have approximately halved. A second prediction of the passive learning model is also borne out in the Belgian data. Surviving young firms exhibit high growth rates in the early years after entry, but growth slows down rapidly with age. The average growth rate declines convexly as it converges to a constant steady state – panel b. of Figure 3.A.1.

As young firms have much higher growth rates and are overrepresented in smaller size classes, the changing composition of the sample leads to a negative relationship between growth and size in a cross-section of firms if we pool all ages. Such a relationship has often been documented in the literature and it is also what we find for Belgium, as shown by the 'all firms' line in Figure 3.1. Growth rates among all firms that survive from year *t*-1 to *t* decline monotonically with the current size of the firm. It is instructive, however, to separately consider the size-growth relationship for young firms of at most six years old, and that of older firms. The dashed line at the bottom of Figure 3.1 shows low growth rates for incumbents regardless of firm size. For them, absolute employment growth is proportional to the current size of the firm, confirming an empirical regularity found in many previous studies.[23] In contrast, growth rates for young firms are not only higher, they clearly increase with size.

Figure 3.1 Growth rates of surviving firms by size: young firms versus incumbents



Note: Annual averages over the 2003-2012 period. We use the dynamic size classification as benchmark method to construct the X-axis. For young firms the 32-63 size class is really 32+, but very few *de novo* entrants have more than 63 employees (shown below).

[23] The second proposition of Gibrat's Law is not confirmed in our dataset. Table 3.A.4 in the Appendix shows that standard deviations of growth rates for mature firms are decreasing in firm size.

The patterns described so far are in line with results from other studies based on large-scale firm-level datasets, even when no or little attempt has been made to distinguish between what we have labelled *de novo* and spurious entrants. It suggests that most patterns are fairly robust to less accurate identification of truly new and young firms. When calculated using the raw administrative data, we indeed find almost the same results for incumbents and all firms as in Figure 3.1.[24] The positive relationship between growth and size that we observe among young *de novo* firms, however, is not replicated in the raw sample of administrative entrants. Instead, the light gray line in Figure 3.1 for the unadjusted administrative data suggests that among young firms, small firms have higher growth rates than larger ones. In Section 3.5.3 below, we show that the difference between *de novo* and administrative entrants is even more pronounced when growth rates are estimated conditional on age, and how spurious entry biases the estimated relationship.

As discussed before, the passive learning model of Jovanovic (1982) has no prediction on the size-growth relationship for young firms. Only with some functional form restrictions does it predict growth to be independent of firm size for all age cohorts. Whatever form the relationship takes, as long as small size classes have relatively more young firms and surviving young firms have higher growth rates—two confirmed predictions of the model—the size-growth relationship is guaranteed to be a negative in the full population of firms. At least if the composition effect is strong enough to overturn the positive relationship for young firms. This is certainly the case in our sample of *de novo* entrants, a finding we turn to first in the next section.

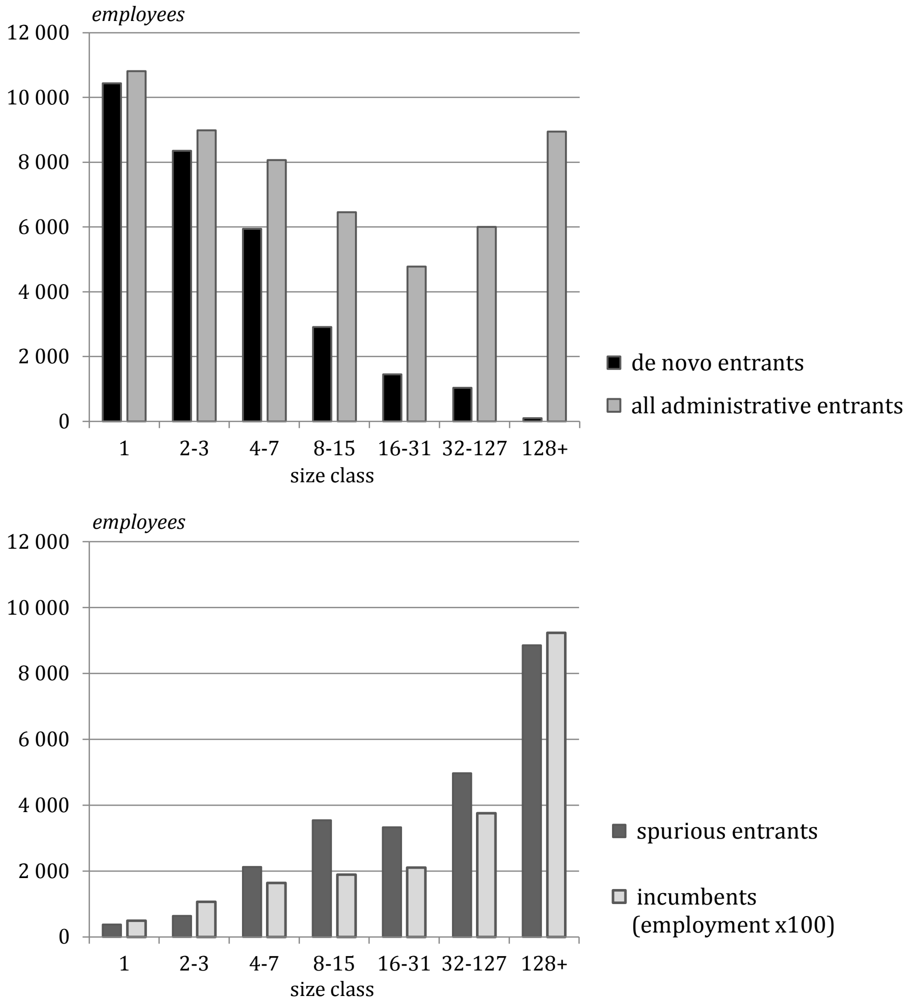
3.5.2 Entry distribution

Although summary statistics based on all administrative entrants or limited to the set of *de novo* entrants look very much alike, a closer examination of both samples reveals some fundamental differences. This is because spurious entrants—pre-existing firms that underwent some reorganization and are misclassified as entrants—introduce incumbent-like features into the population of administrative entrants. As a small group they have little impact on average statistics, but they strongly affect the entry distribution by size or the size-growth

[24] The results calculated using the raw administrative data are reported in Geurts and Van Biesebroeck (2014).

pattern, especially if we use weights to reflect the aggregate employment evolution.

Figure 3.2 Employment distribution of entrants



Note: Annual averages over the 2003-2012 period.

The importance of identifying entrants correctly is readily seen from the employment distribution at entry by firm size. Figure 3.2 shows average annual employment divided into seven size classes on a logarithmic scale. The upper panel shows the employment distribution of *de novo* entrants (dark) against that of all administrative entrants (light). It is well-known that new firms predominantly enter in the smaller size classes, but the distribution based on the administrative sample greatly understates this pattern. Employment of *de novo* entrants is almost entirely concentrated in the first three size categories, which account for fully 82 percent of total job creation of new start-ups. Firms entering with at least 32 employees are exceedingly rare and account for less than 5 percent of total job creation.

The distribution of spurious entrants—the difference between the two series in the left graph—mirrors this pattern. It is mainly concentrated in the larger size classes. The lower panel shows the employment distribution of spurious entrants (dark) relative to that of incumbents (light). The cumulative employment share of the first three size classes is only 13 percent for spurious entrants, while firms with at least 32 workers employ 58 percent of the group's total. The employment distribution of spurious entrants is remarkably similar to that of incumbents. It confirms that spurious entrants are a subset of older firms and suggests that their incidence is unrelated to firm size.

As we have shown, the sample of administrative entrants that uses untreated firm-data mixes two distinct populations of firms. Failing to distinguish between them, as is generally not done, has two implications. First, the size distribution of entrants has a much more dispersed shape than the strong right-skew we observe for *de novo* start-ups. Second, given that employment by spurious entrants accounts for 44 percent of the total in the sample of administrative entrants, it gives an inflated impression of the importance of new firms for job creation in official statistics. In an average year, new job creation by all *de novo* entrants only represents 1.5 percent of the Belgian private-sector workforce. Using administrative entrants instead would suggest this fraction is 2.7 percent, 1.8 times higher.

Besides eliminating false entrants with incumbent-like characteristics, our focus on *de novo* entrants has another important implication. It shrinks the firm sizes that we observe for entrants to a very narrow range. Note that the bottom five size classes, which capture almost all employment of new entrants, are all firms with fewer than 32 employees. This empirical observation is very much in line with the passive learning model, where entrants—having no prior knowledge about their own efficiency—are assumed to all enter at the same size. This is approximately what we observe, and contrasts with the much wider range observed in most previous studies.

The limited size differences we do observe among *de novo* entrants are plausibly the result of selection and growth effects occurring between a firm's startup and the first time we observe it, as new firms only enter the dataset on

June 30. Alternatively, they can reflect some prior knowledge that entrants have about their own intrinsic quality even before they enter the market. The narrow range of observed sizes then implies that a lot is still unknown to these firms when they enter.

It can be expected that spurious entrants also exhibit incumbent-like dynamics following entry and that their overrepresentation in large size classes creates a bias in the size-growth and size-exit pattern for entrants. The bias is hardly noticeable in exit probabilities by size, since exit rates are decreasing in size both for young and older firms. The bias is, however, large in growth estimates by size, where young and older firms strongly differ. This is the topic we turn to next.

3.5.3 Post-entry growth

Most previous studies that empirically examined the relationship between growth and size of young firms have taken for granted firm entry, exit and growth as observed in the data, or applied only a rough correction for spurious entry and exit.[25] It is thus unlikely that reported empirical patterns refer to a well-defined set of truly young firms. Including spurious entrants does not markedly affect many entry and post-entry patterns, as illustrated above. It does, however, bias the size-growth relationship of young firms. Only Haltiwanger et al. (2013) use a dataset which has been edited by advanced record linking methods to distinguish between real and spurious entry and exit.[26] It is therefore not surprising that our results are more in line with that study.

[25] The problem that large-scale firm-level data suffer from spurious entry and exit due to administrative or legal changes, has been recognized since the nineties. However, only with the recent development of advanced record linkage methods, has the extent of the problem and its profound impact on empirical results become clear. Most previous studies did not or could not address this problem. Evans (1987a, 1987b) uses U.S. data from the Dunn and Bradstreet files which are known to suffer from data problems with respect to young and small firms (Davis et al. 1996). Almus and Nerlinger (2000), Lotti et al. (2003) and Mata (2004) do not report the use of linkage methods to clean the sample from spurious entry and exit. Wagner (1994) recognizes that large entry is unlikely and therefore excludes the largest firms from the entry sample, ignoring that spurious entrants also occur in other size classes. Dunne et al. (1989), using the U.S. Census of Manufacturers, partially correct for ownership changes but not for other administrative changes or changes in firm structure.

[26] Unfortunately, most studies do not report the employment distribution at entry, which would be informative about the size range and employment share of larger entrants. Haltiwanger et al. (2013) use traditional record linking methods to eliminate spurious entrants, but additionally rely on physical addresses to more accurately identify entry and exit of multi-establishment firms. It is unclear to what extent their approach identifies all

A positive relationship between growth and firm size

We show the size-growth relationship of *de novo* firms in their first years after entry in Figure 3.3. We then illustrate the robustness of the pattern in Figure 3.4 and describe how delayed adjustment can explain it. In Appendix 3.E, we discuss likely reasons why previous studies did not find the same pattern.

Figure 3.3 plots the coefficients from the employment growth regression of *de novo* entrants that survive from period t-1 to t. Due to the weighting, they represent the net employment growth rates of the entire group of survivors within each age-size class. The benchmark results use the dynamic size classification to assign firms to a size class, while results using two alternative classification methods follow below. As discussed, each method represents an alternative way to classify firms by current size to approximate a continuous size-growth relationship. For clarity, we do not show confidence bounds but report all coefficient estimates and standard errors in Table 3.A.3 in the Appendix. Coefficients are estimated extremely precisely and almost all successive point estimates are significantly different.

As can be seen from the ordering of the different curves, growth rates decrease with firm age when firm size is held constant, in line with the prediction of the passive learning model. In the first year after entry (age 2), surviving young firms of all sizes exhibit very high growth rates. Thereafter, growth rates decline monotonically with age within every size category. Growth rates fall most strongly between age 2 and age 3, and decline at a decreasing rate when an entry cohort matures. Incumbents (labeled age 6+) exhibit growth rates close to zero in all size classes. The convergence of young firm's growth rates to the pattern for incumbents has not been completed entirely when entrants reach age 6, i.e. when we have observed them for five years.

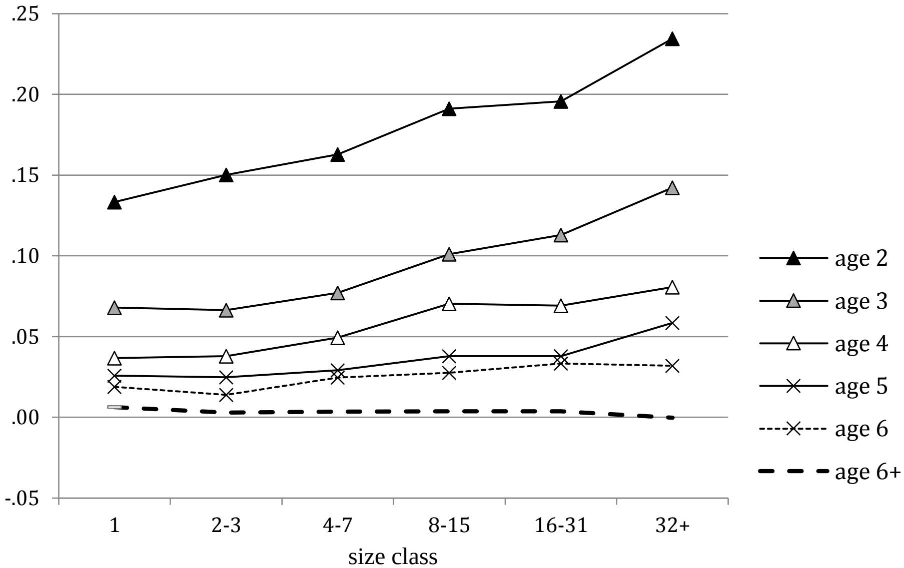
The more remarkable pattern in Figure 3.3 is that growth rates are strongly increasing in current size for firms of the same age cohort. Larger firms grow on average more rapidly than smaller firms of the same age. The positive relationship between growth and size is most pronounced in the first year after entry and gradually weakens with age. Already at age 6, five years after entering the dataset,

---

spurious entrants, especially in medium and large size classes where we relied heavily on the employee-flow method. In their sample, larger firms still represent an important share of employment at entry. Firms entering with more than, respectively, 20 or 250 employees represent 50% or 18% of employment at entry. The corresponding shares in our sample are only 7% or less than 1%. At a minimum, it is likely that their entrant population is not limited to *de novo* entrants as we defined them.

the relationship has shifted towards growth rates that are almost proportional to the current size of the firm. The point estimates for incumbents suggest that growth rates will continue to decline and eventually converge to growth rates close to zero in all size classes. This contrasts with the exit probabilities, which are inversely related to size even for older cohorts.[27] For the smallest firms, growth has basically stalled after five years while for larger firms growth will remain positive for a few years longer. As a result, the firm distribution will continue to shift to the right as illustrated in Figure 3.A.3 in the Appendix.

Figure 3.3 Growth rates of surviving *de novo* entrants by age and size



Note: Annual averages over the 2003-2012 period.

[27] In the passive learning model, the persistent negative relationship between exit rates and size among mature firms is explained by the dependence of the firm's value function on realized costs which are subject to random transitory shocks. Firms terminate their activities when they perceive adverse changes in the distribution of their future profits. As the firm ages, the difference between expected future profit and current profit diminishes because of the increased precision of the firm's information about its own efficiency. However, the firm's decision to stay in the market is based on its realized costs which also depend on firm-specific stochastic shocks that vary from time to time. The firm's value of continuing in operation in the next period is determined by the joint distribution of realized costs in all past periods, hence the dependence on past realizations does not erode away as time progresses. Firms that received negative cost shocks in the past will be smaller at all ages than equally efficient firms that received positive shocks. This induces exit of the smallest marginal firms.

Figure 3.A.4 and Table 3.A.4 in the Appendix provide information of the growth rate distributions of surviving firms in each age-size category. Figure 3.A.4 shows the distributions at age 2 and Table 3.A.4 presents summary statistics at all ages. Growth rates in all subsets tend to be similar. They show a rather symmetric shape which deviates moderately from a normal distribution in two ways. First, growth rates are relatively concentrated in the middle of the distribution. At age 2, about 75 percent of the firms exhibit growth rates that deviate less than one standard deviation from the mean. As a cohort matures, growth rates further tend towards the mean, with about 80 percent of the firms older than age 6 exhibiting growth rates smaller than one standard deviation. This pattern is consistent in all size classes, although the variation in growth rates tends to decrease more strongly among larger firms. As a result, standard deviations are decreasing in firm size at age 6, as they are for more mature firms. The second way in which variation in growth rates among young firms differs from a normal distribution regards its skewness. Growth distributions at early ages are moderately right-skewed and move towards a symmetric shape when age increases. This pattern is highly comparable across size classes. In the first year after entry, about 9 percent of the firms are in the left tail of the growth distribution (smaller than one standard deviation below the mean), and about 16 percent in the right tail. Firms older than 6 years are equally distributed in the left and right tails, and the mean and median growth rates have almost converged.

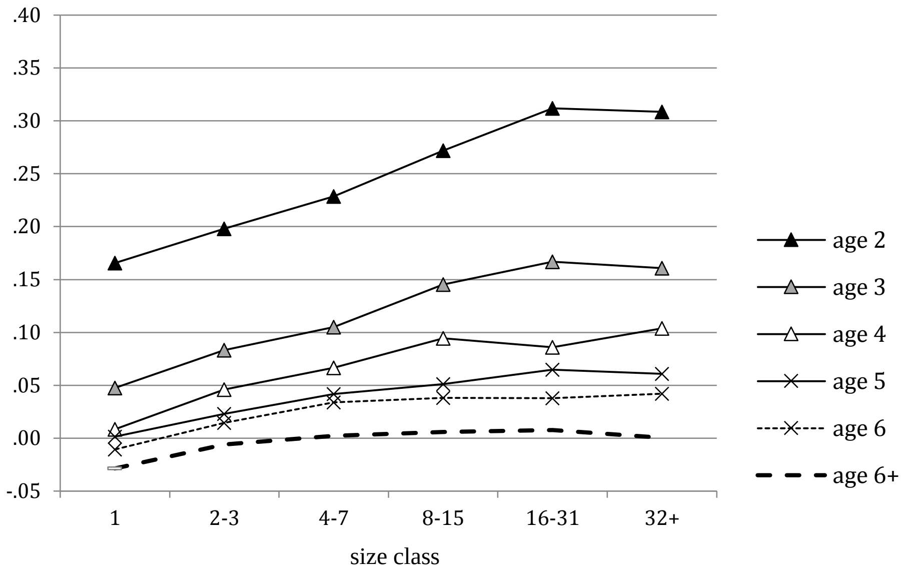
A robust relationship

We conduct two robustness checks that confirm the positive relationship between growth and size of *de novo* entrants. The first presents results based on alternative size classifications. The second verifies whether the positive relationship between growth and size holds in all sectors.

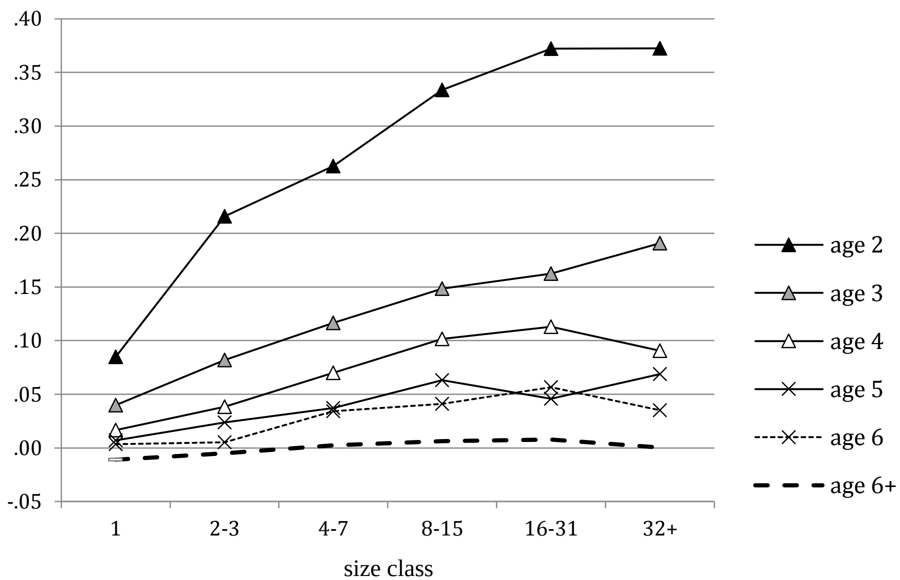
As most *de novo* entrants start with very few employees, we measure firm growth in the following years over a much narrower range of small size classes than is usually the case in other studies. This heightens the statistical problems associated with the conventional base-year size classification that we discussed earlier. To complement the results based on the dynamic size classification, we show in Figure 3.4 estimates based on two alternative size classifications. Panel a. presents results that average over growth rates using the beginning-of-period and end-of-period sizes as base. In panel b. firms are classified by the average of their size in years *t*-1 and *t*, as in Haltiwanger et al. (2013).

Figure 3.4 Alternative size classifications: growth rates of surviving *de novo* entrants

a. Average of estimates using firm base size at *t*-1 and *t*



b. Average size classification



Note: Annual averages over the 2003-2012 period.

The patterns using both alternative methods are similar to the benchmark results. Growth rates are increasing in firm size within each age class. The strong positive slope in the first few years following entry gradually converges to a virtually flat profile for incumbents. The positive relationship is somewhat more pronounced than in our benchmark results, especially in panel b., where job gains of fast-growing firms are entirely allocated to the intermediate size class between t-1 and t. In the dynamic size classification, this growth is allocated to each respective size class the firm passes through.

It is quite remarkable that across the three graphs, there is only a single instance where any of the curves intersect. The patterns we uncover are very smooth and monotonic: growth rates increase with size for each age cohort and decrease with age for each size class. This is even more remarkable given that they have been estimated over the very turbulent 2003-2012 period that includes the Great Recession. The patterns also hold if we limit the sample to firms entering between 2003 and 2007 and follow their growth to at most 2008, the onset of the crisis, or if we limit the sample to firms entering from 2008 onwards.[28]

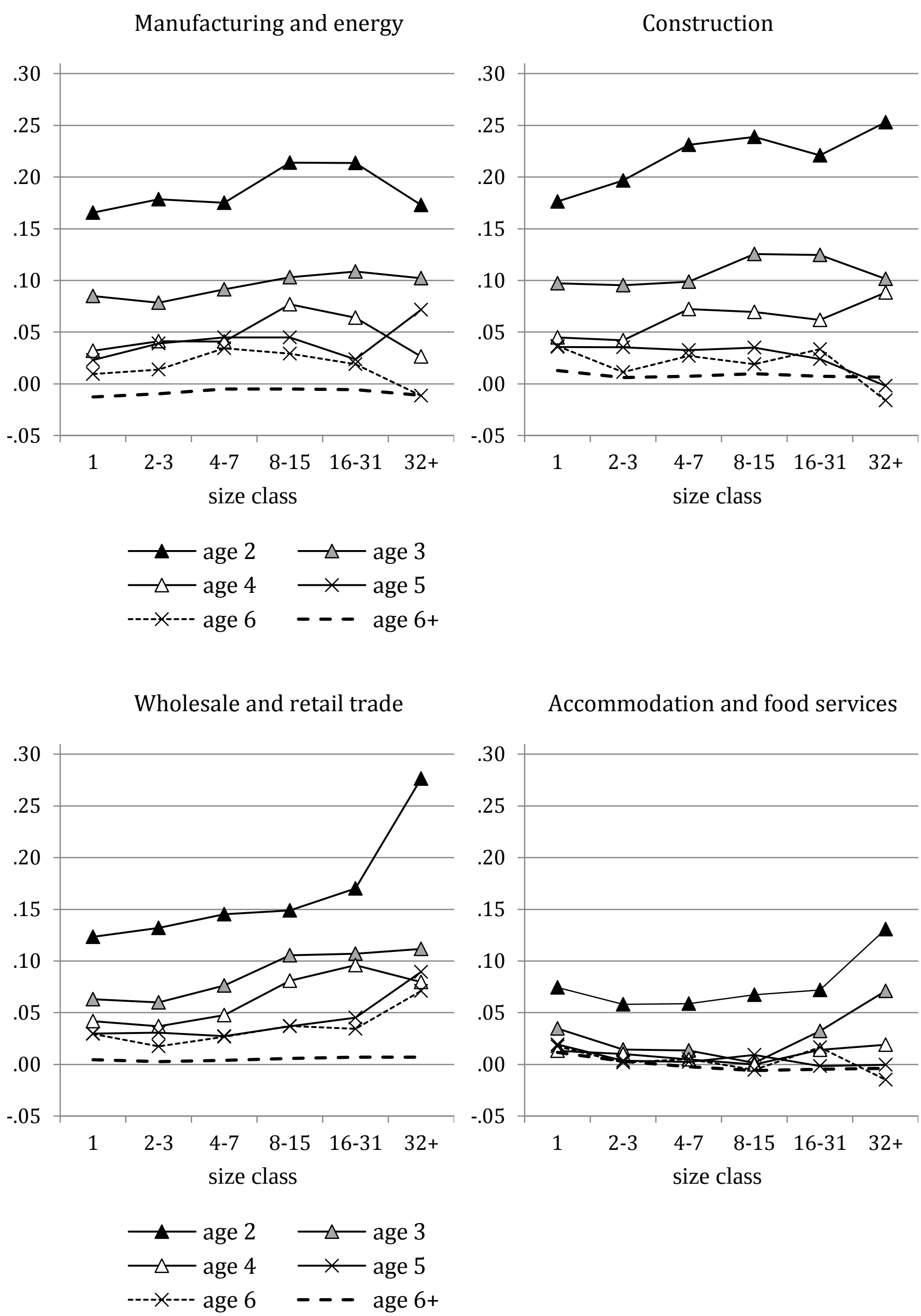
The positive relationship between growth and size of young firms of the same age confirms the results in Haltiwanger et al. (2013) that are obtained using the average size classification. As noted before, that study uses a dataset and size measurement that reduce potential biases when working with small and young firms. An important difference, however, is that the growth rates they report do not evolve to size-invariant growth among older firms, while many studies have found that proportionate growth rates a good approximation of the size-growth relationship among large and well-established firms (Mansfield 1962; Hall 1987; Geroski 1995). In addition, a positive size-growth pattern cannot be a steady state as the firm size distribution would collapse.

Figure 3.5 presents the relationship between growth and size of *de novo* entrants for six broad industry groups: Manufacturing, Construction, Trade, Accommodation & Food Services, Business Services, and Mixed Household & Business Services. Growth rates of surviving *de novo* entrants in all industries show broadly the same pattern as in Figure 3.3. 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

[28] Separate results for pre and post-crisis entrants are shown in Geurts and Van Biesebroeck (2014).

diversification after entry resulting in little difference between growth rates from age 3 onwards.

Figure 3.5 Growth rates of surviving *de novo* entrants in six industry groups



(Figure 3.5 continued)

Note: Annual averages over the 2003-2012 period.

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 is 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. The increasing pattern is least pronounced in Manufacturing. 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 16 employees in most age cohorts. The results for this sector thus do not differ entirely from the three previous studies that found a negative growth size relationship and are each based on samples of young firms in Manufacturing only (Evans 1987a; Lotti, Santarelli and Vivarelli 2003; Mata 1994).

What explains the positive relationship?

We have argued that the positive relation between growth and size among young firms of the same age cohort is not at odds with the predictions of Jovanovic (1982) if one takes into account that young firms exhibit some lag of adjustment to prior information.[29] Firms that receive positive information, i.e. learn that they are more efficient than previously realized, will not always adjust completely to this new information right away. Risk aversion might induce them to wait an extra period for the positive information to be confirmed or it might take some time for additional capacity to become operational. Financial constraints, hiring frictions, or regulations can impose external barriers that need to be overcome before a firm can expand its operations. Such partially delayed growth will induce a positive size-growth relationship. Some of the positive news leads to instantaneous growth and raises a firm's current size. The remaining fraction of growth postponed to subsequent years then leads to a positive correlation between growth and size.

A corresponding delay for firms that adjust to negative information will further strengthen the positive correlation. If annually recurring fixed costs of operation are sufficiently low relative to sunk entry costs, firms might delay their eventually withdrawal from the market even as they make losses. In the data we even observe many firms with no employees for some years. It suggests that merely surviving might not be all that costly. As firms adjust their size downward but postpone exit, it leads to low or negative growth rates for smaller firms.

Figure 3.A.5 in the Appendix provides some evidence for such behavior. In panel a., firms that are about to exit in the next period exhibit much lower growth rates than firms that will survive. The difference is approximately constant in each of the 5 years following entry. Average growth rates are negative for impending exiters at all ages except age 2, indicating that firms stay small or decline in the year before they exit. The difference in growth rates already appears two years before exit, shown in grey, but is less pronounced. Given that there are many more firms exiting in the smaller size classes, this pre-exit growth difference contributes to the positive size-growth relationship.[30]

[29] While the model in general has no prediction for the size-growth relationship conditional on age, under some assumptions—in particular constant returns to scale—growth rates should be size invariant.

[30] Panel b. in the same figure shows that delayed exit does not explain the observed positive size-growth relationship entirely. Excluding all *de novo* entrants that exit before age 6, growth rates still show a positive relationship in the first years that gradually convergences to a size-invariant pattern.

3.6 Conclusion

In constructing the dataset, we have taken great care to identify a sample of firms that start new operations, corresponding to actual new firm creation. Complementing the traditional firm linkage method with an employee-flow method, we filtered out misclassified, spurious entrants. Given their incumbent-like entry distribution and growth patterns, they bias the patterns of interest. By establishing a more complete set of firm linkages, we also avoid confusing firm restructuring events with economic exits. For the remaining group of *de novo* entrants, we confirm several patterns from the literature. In particular, exit rates are shown to be strongly declining in age and size, while growth rates for survivors decline with age and also with size if we pool across age cohorts.

In addition, we obtain two novel findings. First, we find that firm entry sizes are reduced to a narrow range of small size classes. Second, growth rates of *de novo* entrants are increasing with size in the first years, but quickly converge to proportionate growth as an entry cohort matures. The firm size distribution at entry differs more markedly from that of mature firms than is usually the case, but the positive size-growth pattern accelerates the tendency towards increased concentration in an entry cohort and leads to a pronounced right shift in the firm size distribution.

The exit and growth patterns by age and size class are remarkably regular. We have estimated them over an extremely turbulent time period that includes the Great Recession, but all age and size patterns are entirely monotonic. The persistent features of firm dynamics of very young firms seem to dominate cyclical factors.

Our results are consistent with firms having a very imperfect knowledge of their productivity at entry. All patterns are in line with the passive learning model of Jovanovic (1982) where a firm's underlying efficiency is constant, but is only discovered as a firm operates in the market. If we add delayed adjustment, both in exit and in growth, even the positive size-growth relationship for young firms is consistent with the model.

Note that frictions could even influence firms' choice of initial entry size. Evans and Jovanovic (1989) provide an alternative model where liquidity constraints force some firms to enter below their desired size and grow into their optimal size afterwards using retained earnings to expand. This would lead to a negative size-growth relationship as constrained, smaller entrants would have a greater upside potential. Both the narrow firm size distribution and the positive size-growth relationship we have documented are more supportive of constraints affecting firms following their entry decision rather than before.

Our findings suggest some cautious policy conclusions. A recent literature has documented that especially in less developed economies, production factors are often stuck at unproductive firms (Hsieh and Klenow 2009). This type of misallocation lowers potential output and aggregate productivity. Our evidence suggests that new firms do not know their own likelihood of success very well and it is inevitable that some unproductive entrants end up with too much resources. Policy can accommodate this by making sure that adjustments to firm size after entry are easy to make. At the same time, lowering entry barriers in a situation where adjustment frictions after entry are large is likely to generate bad aggregate outcomes.

It is, however, not straightforward to draw strong policy inferences from the empirical regularities presented for *de novo* entrants. As demonstrated by Brown et al. (2015), subsets of firms which exhibit the highest growth rates are not necessarily the ones that experience the strongest constraints on growth and would respond the most to policy intervention. In particular, they show that in a cross-section of firms of all ages, the job creation effect from loan programs increases with firm size, while employment growth rates in the same population are negatively related to firm size. They do find evidence, however, that fast-growing firms are the ones that experience the greatest constraints to growth.

We have suggested that delayed adjustment is one mechanism that can explain the observed positive size-growth relationship. In increasingly global markets and with rapid technological advancement, such growth delays can be quite costly. New entrants often have only a narrow window of opportunity to occupy a market niche. If scaling-up in response to positive information happens too slowly, a firm risks coming too late and be shut out of the market by early movers.

Guner et al. (2008) provide evidence that many government policies favor small firms. This is often rationalized on the assumption that small firms are the engine of job creation in the economy. Previous literature has already highlighted that one should not confuse the (conditional) effects of age and size—it tends to be young firms which are vital for job creation. Our current findings cast further doubt on the employment growth potential of small entrants. Among young firms of the same age, those showing up in the dataset with a smaller size also tend to grow more slowly subsequently.

In continuous time, one can think of firm entry as the moment the first employee is hired. Some entrants add additional employees in the next minutes or days, while others take years. With adjustment frictions, it is likely that a size-pattern established early on will be perpetuated over time. A small size, conditional on age, is indicative of negative news about a firm's profitability early on. While not all firms can freely choose their size—a large literature documents constraints and frictions that limit a firm's initial size—our overall patterns suggest that by and large small firms choose to be small. Directing subsidies primarily towards the smallest firms or imposing size restrictions to qualify for government support are policies that should be avoided.

Appendix

3.A Tables and Figures

Table 3.A. 1 Employee-flow linkages by decision rule

a. Type of employee-flow linkages by decision rules

See Table 2.A.5 in the Appendix of Chapter 2

b. Share of employee-flow linkages by type

	All links	Spurious entrants	Transfers
1. ID-change (largely identical)	0.57	0.78	0.71
2. Takeover (absorption) 75%	0.22	-	0.15
3. Split-off 75%	0.12	0.18	0.05
4. Takeover (absorption) 50%	0.01	-	0.02
5. Split-off 50%	0.01	0.01	0.01
6. Merger of exits	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.

Table 3.A. 2 Summary statistics for *de novo* entrants

	Entry rate	Employment share	Exit rate	Share of survivors	Employment share of survivors	Average size (employees)
Age 1 (entry)	0.09	0.015		1.00	1.00	1.93
Age 2			0.21	0.79	0.98	2.39
Age 3			0.15	0.68	0.98	2.78
Age 4			0.13	0.60	0.98	3.10
Age 5			0.11	0.54	0.98	3.38
Age 6			0.10	0.49	0.98	3.61

Note: Annual averages over the 2003-2012 period. The year a firm enters the dataset is indicated by age 1.

Table 3.A. 3 Growth rates of surviving *de novo* entrants by age and size
Coefficient estimates shown in Figures 3.3 and 3.4. Standard errors in parentheses.

a. Dynamic size classification

	Firm size class (employment)					
	1	2-4	4-7	8-15	16-31	32+
Age 2	0.133	0.150	0.163	0.191	0.196	0.234
	(.002)	(.002)	(.002)	(.003)	(.004)	(.005)
Age 3	0.068	0.067	0.077	0.101	0.113	0.142
	(.003)	(.002)	(.002)	(.003)	(.004)	(.004)
Age 4	0.037	0.038	0.049	0.070	0.069	0.081
	(.004)	(.003)	(.003)	(.003)	(.004)	(.004)
Age 5	0.026	0.025	0.029	0.038	0.038	0.059
	(.004)	(.003)	(.003)	(.003)	(.004)	(.004)
Age 6	0.019	0.014	0.025	0.028	0.033	0.032
	(.005)	(.003)	(.003)	(.003)	(.004)	(.004)
Incumbents	0.006	0.003	0.004	0.004	0.004	0.000
(age 6+)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)

b. Average of estimates using firm base size at *t*-1 and *t*

	Firm size class (employment)					
	1	2-4	4-7	8-15	16-31	32+
Age 2	0.166	0.198	0.229	0.272	0.312	0.309
	(.004)	(.004)	(.004)	(.006)	(.007)	(.008)
Age 3	0.048	0.083	0.105	0.145	0.167	0.161
	(.004)	(.003)	(.003)	(.004)	(.005)	(.005)
Age 4	0.009	0.046	0.066	0.094	0.086	0.104
	(.005)	(.004)	(.003)	(.004)	(.005)	(.005)
Age 5	0.002	0.023	0.042	0.051	0.065	0.061
	(.005)	(.004)	(.003)	(.004)	(.005)	(.004)
Age 6	-0.011	0.015	0.034	0.038	0.038	0.042
	(.006)	(.004)	(.003)	(.004)	(.005)	(.005)
Incumbents	-0.029	-0.006	0.002	0.006	0.008	0.001
(age 6+)	(.001)	(.001)	(.001)	(.001)	(.001)	(.000)

c. Average size classification

	Firm size class (employment)					
	1	2-4	4-7	8-15	16-31	32+
Age 2	0.085	0.216	0.263	0.334	0.372	0.373
	(.004)	(.003)	(.004)	(.005)	(.007)	(.008)
Age 3	0.040	0.082	0.117	0.148	0.162	0.191
	(.004)	(.004)	(.004)	(.005)	(.006)	(.006)
Age 4	0.016	0.038	0.070	0.102	0.113	0.091
	(.005)	(.004)	(.004)	(.005)	(.006)	(.006)
Age 5	0.007	0.024	0.037	0.063	0.046	0.069
	(.006)	(.004)	(.004)	(.005)	(.006)	(.005)
Age 6	0.003	0.005	0.034	0.041	0.056	0.035
	(.007)	(.005)	(.004)	(.005)	(.006)	(.006)
Incumbents	-0.011	-0.005	0.002	0.006	0.008	0.000
(age 6+)	(.001)	(.001)	(.001)	(.001)	(.001)	(.000)

Table 3.A. 4 Summary statistics of growth distributions of surviving *de novo* entrants

	Firm size class (employment)					
	1	2-3	4-7	8-15	16-31	32+
Mean						
Age 2	.063	.205	.246	.286	.337	.379
Age 3	.029	.077	.109	.135	.166	.178
Age 4	.015	.037	.067	.102	.107	.071
Age 5	.006	.023	.035	.064	.056	.076
Age 6	.005	.006	.032	.040	.053	.045
Incumbents	-.008	-.003	.003	.007	.008	.004
Standard deviation						
Age 2	.310	.540	.524	.570	.625	.592
Age 3	.323	.458	.414	.416	.425	.375
Age 4	.327	.426	.388	.365	.354	.316
Age 5	.321	.401	.353	.317	.307	.280
Age 6	.319	.392	.344	.309	.265	.210
Incumbents	.304	.347	.287	.234	.195	.147
Median						
Age 2	.000	.000	.182	.222	.240	.243
Age 3	.000	.000	.000	.105	.125	.110
Age 4	.000	.000	.000	.087	.080	.048
Age 5	.000	.000	.000	.000	.057	.048
Age 6	.000	.000	.000	.000	.036	.023
Incumbents	.000	.000	.000	.000	.000	.000
Share of firms in left tail (smaller than mean - 1 standard deviation)						
Age 2	.065	.122	.100	.110	.092	.067
Age 3	.097	.146	.116	.106	.099	.072
Age 4	.110	.155	.122	.095	.103	.086
Age 5	.111	.150	.122	.104	.098	.055
Age 6	.111	.149	.110	.098	.100	.107
Incumbents	.108	.131	.104	.093	.078	.070
Share of firms in center (within 1 standard deviation of the mean)						
Age 2	.775	.695	.753	.741	.751	.789
Age 3	.763	.730	.763	.768	.779	.778
Age 4	.759	.755	.725	.782	.781	.786
Age 5	.768	.773	.722	.769	.781	.848
Age 6	.771	.686	.746	.789	.789	.783
Incumbents	.793	.741	.800	.804	.834	.852
Share of firms in right tail (larger than mean + 1 standard deviation)						
Age 2	.160	.183	.146	.148	.157	.144
Age 3	.141	.124	.122	.126	.121	.151
Age 4	.131	.090	.152	.124	.116	.128
Age 5	.121	.077	.157	.127	.121	.097
Age 6	.118	.165	.144	.114	.111	.111
Incumbents	.099	.128	.096	.103	.088	.079

Note: Results based on pooled sample across years (2003-2012) and all industries using the average size classification. Incumbents are firms older than age 6.

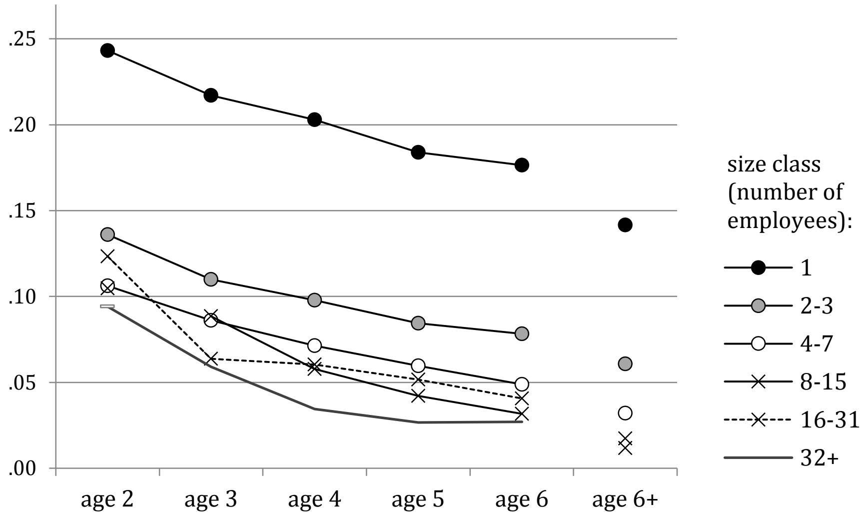
Table 3.A. 5 Six main industries and NACE Rev. 2 classes

Nace Rev. 2 classes	*De novo* entrants: Number of firms	Number of employees	Average entry size (employees)
1. Manufacturing and energy Section B, C, D, E	777	1 996	2.6
2. Construction Section F	2 730	5 150	1.9
3. Wholesale and retail trade Section G	4 236	7 497	1.8
4. Accommodation and food services Section I	2 793	6 570	2.4
5. Business support services Nace 49.2, 49.4, 49.5, 50.2, 50.4, 51.2, 52.1, 52.241, 52.249, 62, 63, 64.110, 64.2, 64.3, 64.910, 64.991, 64.992, 64.999, 66, 69.2, 70, 71, 73, 74, 77.1, 77.3, 77.4, 80, 81 (excluding 81.210, 81.220), 82, 95.1	2 945	5 143	1.7
6. Mixed business & household services Nace 49.1, 49.3, 50.1, 50.3, 51.1, 52.210, 52.220, 52.230, 52.290, 53, 58, 59, 60, 61, 64.19, 64.921, 64.92, 65, 69.1,72, 75, 77.2, 79, 95.2, 96, and Section L	2 011	3 459	1.7
Total	15 492	29 815	1.9

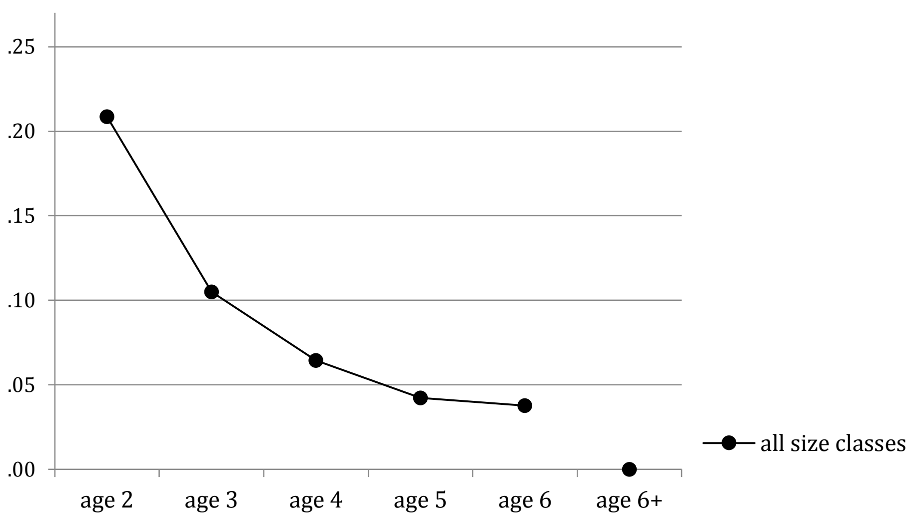
Note: Annual averages (2003-2012) of *de novo* entrants and employment in entry year. Firms not in the listed categories are excluded from the analysis, primarily quasi-public sector services and subsidized household help. The detailed explanation of the Nace codes is provided in Table 2.A.1 of Chapter 2.

Figure 3.A 1 Confirmed predictions of the passive learning model

a. Exit rates of *de novo* entrants by age and size



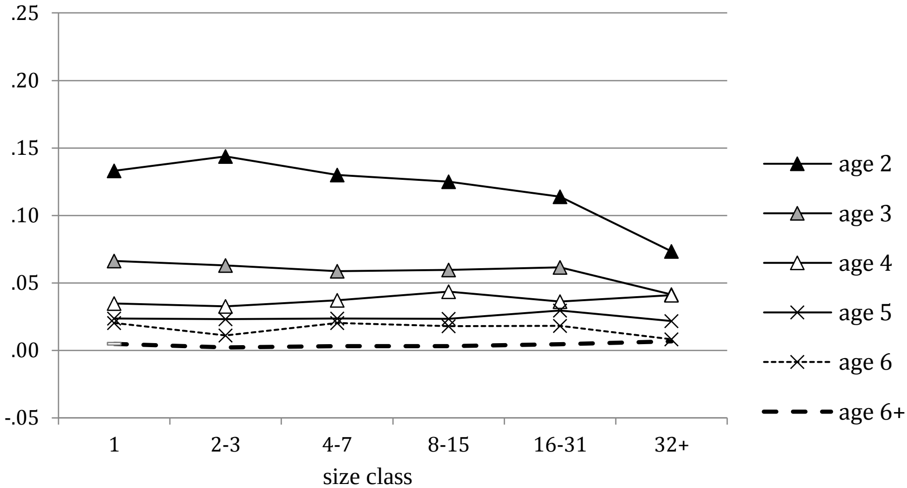
b. Growth rates of surviving *de novo* entrants by age



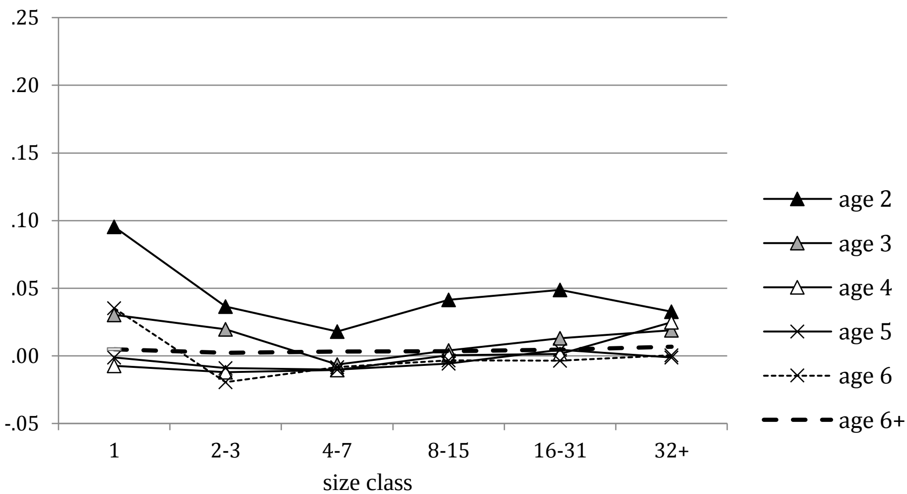
Note: Annual averages over the 2003-2012 period. Age 6+ refers to incumbents.

Figure 3.A 2 Growth rates of administrative and spurious entrants by age and size

a. Growth rates of administrative entrants
(including *de novo* and spurious entrants)



b. Growth rates of spurious entrants



Note: Annual averages over the 2003-2012 period. Age 6+ refers to incumbents.

Figure 3.A 3 Evolution of the firm size distribution

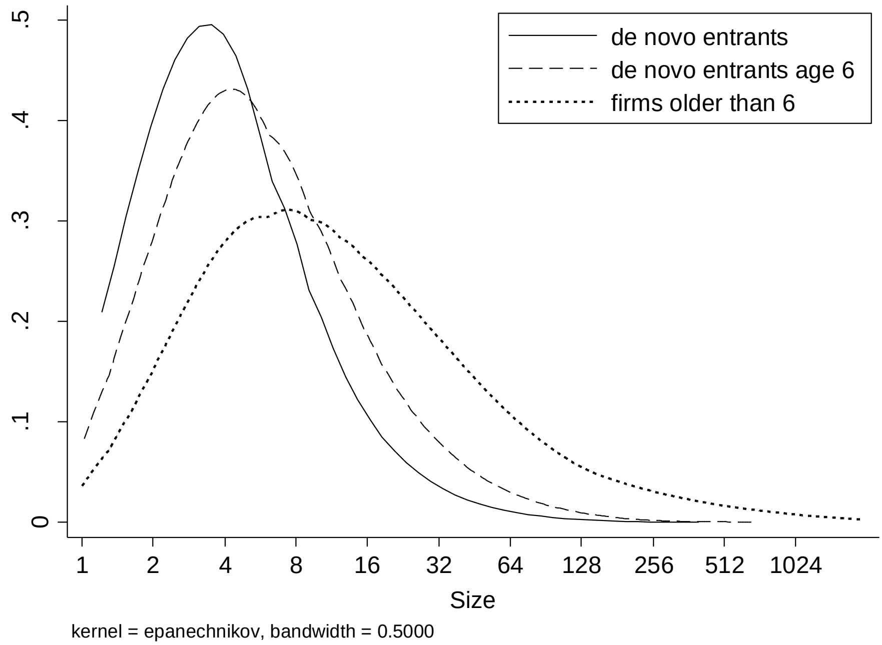
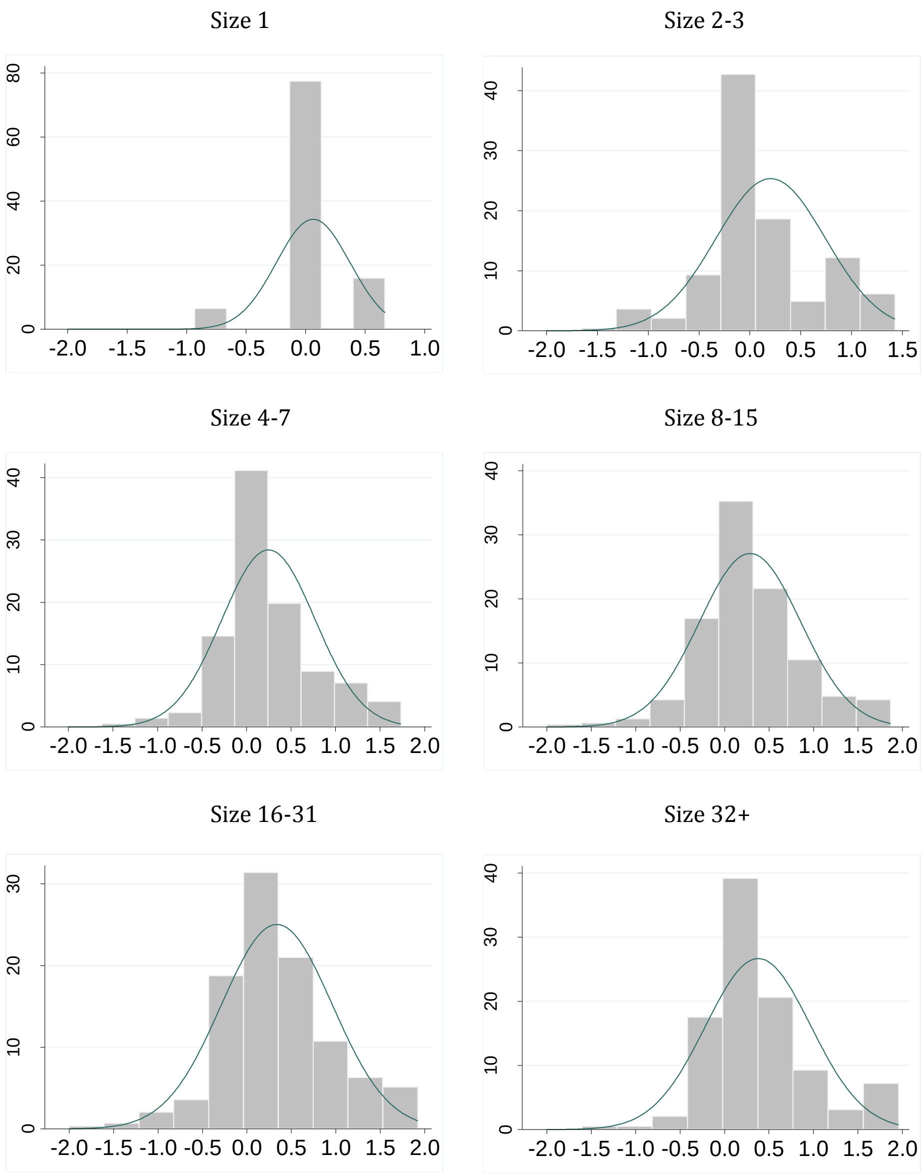


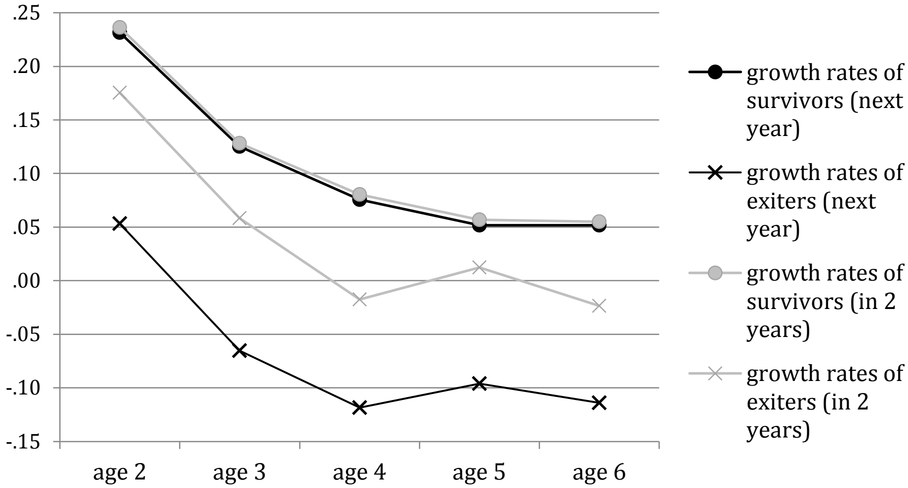
Figure 3.A 4 Distribution of growth rates of surviving *de novo* entrants at age 2 by firm size class



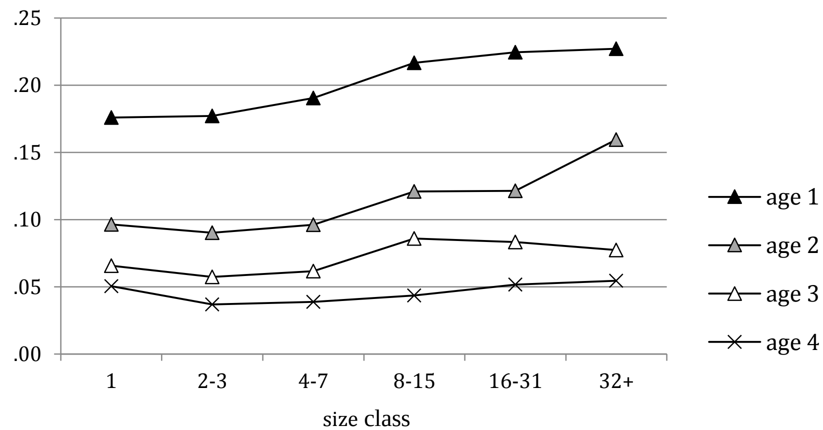
Note: Results based on pooled sample across years (2003-2012) and all industries using the average size classification. The solid lines show the normal distribution.

Figure 3.A 5 Delayed adjustment of *de novo* entrants in exit and growth

a. Delayed exit: growth rates of survivors versus exiters

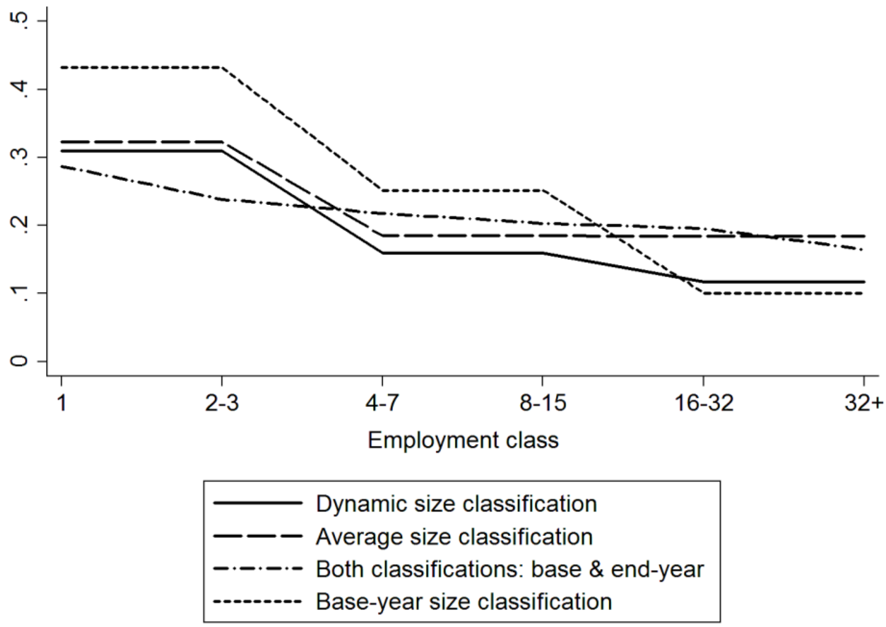


b. Delayed growth: growth rates of firms surviving till age 5



Note: In panel b. we use the dynamic size classification.

Figure 3.A 6 Estimated size-growth relationships on simulated data with constant growth rate



Note: Calculations on simulated dataset with growth rates that are size-invariant.

3.B Data

The analysis is based on a firm-level dataset maintained by the National Social Security Office (NSSO) of Belgium. It covers the universe of firms with at least one employee over the period 2003-2012. For comparability with other studies, we restrict the analysis to firms in the private, non-farm sector and also exclude highly subsidized sectors which receive strong support from government programs.[31] In an average year, the sample includes 178 000 firms and 2 070 000 employees. Total employment increased during the sample period by 0.9 percent per year till 2008, dropped by 2.5 percent between 2008 and 2010 and has been more or less stable since.

Large-scale firm-level data collected for administrative or statistical purposes have become the main information source for empirical analysis on firm dynamics. A drawback of these data, however, is that changes in ID code or firm structure lead to missing linkages in the longitudinal observation of firms. This mistakenly introduces entry and exit events, as well as spurious shocks to firm-level employment growth. We refer to Chapter 1 for a detailed discussion of the problem and the solutions that we have implemented. Here, we only summarize the strategy that has been adopted to identify *de novo* entrants and their post-entry employment histories.

The first linking method we apply has been developed by Statistics Belgium and implements the OECD-Eurostat recommendations on business demography statistics (Eurostat-OECD 2007). It exploits information on firm continuity from a comprehensive database that combines information from different administrations such as the national register of legal entities, the trade register, VAT declarations, and Social Security reports. In addition, it relies on a probabilistic matching procedure that uses similarities in firm name, address, and industry code to link different ID codes of the same firm across two years.

Our second linking method uses a definition of firm continuity that is based on its workforce. It follows one of the main production factors of the firm, the stock of employees, to trace changes in ID codes and firm structure. It exploits the linked employer-employee information in the NSSO dataset: both firms and employees are identified with a unique ID code. The advantage is that an individual never

[31] Table 3.A.5 lists all NACE sectors we include in the analysis and classifies them into six industries. Excluded sectors include "Human health and social work activities," where most expenditures are publicly financed, and "Subsidized household help," where service vouchers subsidize 70% of the wage cost.

changes ID and can always be followed. If a firm changes ID code but continues its activities, the stock of employees will largely be the same for the old and the new firm ID. Similarly, when firms merge or split up, this will be reflected in a merge or division of workforces. Continuity of the workforce can thus be used to identify firms that operate continuously but change ID code or firm structure.

In practice, we follow clusters of employees that move simultaneously from one ID code to another between two quarterly observations. A set of decision rules regarding the size of the employee cluster relative to the firms' total workforce is used to determine whether we should consider the two ID codes as a single, continuing firm. The primary rule, to identify one-to-one ID changes, verifies whether the cluster represents at least 50 percent of the workforce of both the disappearing and the newly appearing ID code. A second rule identifies takeovers, allowing the receiving ID code to exist already, but requiring a cluster of at least 75 percent of the workforce of the initial ID code to move together. A set of additional decision rules is listed in Table 3.A.1 and these capture takeovers, split-offs and other forms of organizational restructurings. The table shows that the first two rules account for 80 percent of the identified links. In Chapter 2, we conduct several robustness checks to verify the sensitivity of measures of firm dynamics to alternative size thresholds and decision rules of the employee-flow method. We find that they are not critical to the empirical results.

The linkages established by the two record linking methods are first used to identify continuing firms that are misclassified as entrants and exits. They are labeled as 'spurious' entrants and exits as opposed to *de novo* entrants and true exits. Panel b. of Table 3.A.1 shows that 78 percent of the spurious entrants we identify are simply incumbents that continue the same activities with a new identification code after a purely administrative or legal change. Another 18 percent are split-offs of another firm.[32] Second, for those firms that are involved in an ID change or restructuring, administratively recorded employment changes from one period to the next do not reflect internal job growth but are but artificially inflated or deflated by the event. Therefore, as a further step in the data editing, employment of these firms is imputed in the years after the event. Our approach is to construct an aggregate event-level that includes all firm ID's

[32] Some administrative entrants are subsidiaries of foreign firms entering the Belgian market and are not *de novo* entrants either. Our linkage methods are unable to identify these FDI entrants. As it is an extremely small group, their presence is unlikely to affect the results. On a reduced sample, covering the 2005-2010 period, we find that they represent fewer than 1 percent of all *de novo* entrants.

interlinked from *t*-1 to *t*. Firm-level employment in *t* and *t*+n is then imputed by assuming the same growth rate for each firm involved in the event. The imputation procedure is extended to the sixth year of existence for *de novo* entrants. [33] For one-to-one ID changes, which represent the vast majority of events, the imputation method simply corresponds to replacing the new by the old ID code. With respect to more complex events, the imputation method treats break-ups and mergers of firms symmetrically and preserves the firm size distribution in the sample. Imputed employment histories more closely reflect actual job creation or destruction at the firm level and allow a more accurate estimate of post-entry exit and growth patterns by size.

The linkage methods similarly divide the group of *de novo* young firms that disappear from the dataset into true economic and spurious exit. The extent of misclassification is somewhat lower than on the entry side, 4 percent of administrative exits are identified as spurious, but the likelihood is again increasing with firm size. In the working paper, see Geurts and Van Biesebroeck (2014), we report those statistics and provide separate summary statistics for all the different groups of entrants and exiting firms.

3.C Size classification

Regression-to-the-mean and sample selection may spuriously introduce a negative relation in estimates of the relationship between growth and size of surviving firms if firms are classified by their size in the base year *t*-1. The extent to which these problems bias actual empirical results, and possible solutions have been extensively debated in the literature, without reaching a unanimous conclusion so far.[34] As discussed before, both problems are exacerbated if growth rates are measured in a population of predominantly small firms, as is the case in our sample of *de novo* entrants. We therefore need to directly address these measurement problems. To avoid bias in the size-growth relationship, we use three alternative firm-size classifications that approximate a continuous size-growth relationship.

---

[33] We also impute employment for mature firms involved in an event to calculate consistent employment growth rates for them, which we use as a comparison for the evolution of *de novo* firms.

[34] For a discussion see for example Hall (1987), Baldwin and Picot (1995), Davis et al. (1996b), Davidsson et al. (1998), and Kirchhoff and Greene (1998).

The first size classification method, and the one we use for our benchmark estimates, allocates employment gains and losses to each respective size class in which the growth or loss occurred. This 'dynamic' sizing 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), and is further discussed in Davidsson, Lindmark and Olofsson (1998) and de Wit and de Kok (2014). Firms are initially assigned to a size class based on employment in t-1, but are re-assigned to a new class when they cross a threshold. The growth from E_{it-1} to the threshold is assigned to the initial class and the remaining growth from the threshold to E_{it} is assigned to the next size class. Growth rates use average employment in the denominator as discussed in Section 3.4 of the main text, but use the intermediate size class thresholds as upper or lower limits. This methodology approximates instantaneous class re-assignment that would be feasible if size and growth were measured in continuous time. We choose the size class thresholds such that they imply symmetric and (almost) equal ranges of potential growth rates within each class between -0.67 and +0.67.[35] This approach mitigates the negative bias in the size-growth relationship caused by regression-to-the-mean because symmetric growth and decline are equally attributed to the same size classes. The problem of left-truncated growth rates in the smallest size classes is also mitigated because the range of growth rates within each size class is symmetric with mean zero. The equal ranges of potential growth rates further imply that no size class is favored when the sample exhibits on average positive (or negative) growth, avoiding the upward size-growth bias of the methodology used by Haltiwanger et al. (2013) discussed below.

The second classification method uses each firm twice in the regression, assigning a weight of one half to each observation. One observation uses the firm's employment level at the beginning of the period both as a base for the growth rate and to determine the size class. The second observation uses the firm's employment at the end of the period for both calculations. Growth rates of firms assigned to the same size class based on E_{it-1} or E_{it} contribute to the regression in a symmetric way as before. Firms assigned to different size classes can show a different size-growth relationship in each instance and both contribute equally to the average pattern identified in the regression. This approach has been proposed

[35] The size thresholds between the size classes]0,2[, [2,4[, [4,8[, [8,16[, [16,32[, and [32,∝[are 2, 4, 8, 16, and 32 for expansion and 1.85, 3.7, 7.4, 15, 31 for contraction. This yields growth ranges of [-0.60,+0.67], [-0.67,+0.67], [-0.67,+0.67], [-0.68,+0.67], [-0.70,+0.67], [-0.68,+0.67], and [-0.67,∝] respectively.

by Prais (1958) to avoid regression-to-the-mean bias and can be motivated similarly as the use of average wage shares in a Solow residual, i.e. as a discrete approximation to the continuous Divisia index of productivity growth (Caves, Christensen and Diewert 1982).

For comparison with the results of Haltiwanger et al. (2013), our last classification method uses the average of firm size in years *t*-1 and *t* as a proxy for the size over the intervening period. This size classification, proposed by Davis et al. (1996a, 1996b), reduces the regression fallacy and the truncation problem. If firm size fluctuates around a stable long-run size, using the average size classification would yield unbiased results. However, in a sample with on average positive growth rates, it introduces an upward bias between size and growth (Baldwin and Picot 1995).[36] Rapidly growing firms are more likely to cross a size class border and their measured rate of growth will be entirely reassigned to a higher size class.

In Figure 3.A.6, we report regression results on a simulated dataset where we imposed the same average growth rate for all size categories. We started from a cohort of *de novo* entrants that replicates the actual entry size distribution observed in the data. We then applied a stochastic growth rate to each observation that averaged 10 percent regardless of size, but with a large dispersion, as in the observed data. We then applied an exit rule that was stochastically decreasing in firm size, generating an exit probability that is negatively correlated with the growth rate. The size-growth relationship was then estimated using each of the size classification methodologies just discussed and also using the base-year classification. The graph plots the regression coefficients on the different size class dummies. The results confirm the strong downward bias in the size-growth relationship for the base-year classification and a much more constant relationship for the three alternatives, especially for firms with at least 4 employees.

3.D Confirmed patterns

As found in many other countries, the annual entry rate is high but involves only a small fraction of the labor force. Statistics in Table 3.A.2 show that *de novo* entrants represent 9 percent of all active employer firms in a given year, but only 1.5 percent of total employment. Most entrants are small. Average entry size is

[36] For further discussion see also Davidsson et al. (1998) and Kirchhoff and Greene (1998).

1.9 employees, six times smaller than the average size of incumbents. In the years following entry, a large fraction of the entering cohort exits and the average firm size among survivors increases. Only half of all entrants are still around at age 6, at which time the average firm size in the surviving group has almost doubled. Job creation by survivors is substantial and almost compensates for job loss due to the exit of young firms. Total employment created by an entry cohort falls only slightly below its initial value in the five years after entry.

As the entry cohort matures, the size distribution becomes more concentrated as illustrated by the kernel density in Figure 3.A.3. The strongly right-skewed distribution at entry gradually gets a fatter right tail, but at age 6 it has not yet converged to the distribution of incumbents.

A first mechanism generating this pattern of increased concentration in an entry cohort is selective survival. In line with the predictions of the passive learning model, we find high exit rates for young firms which are decreasing in age as well as in size. This is shown in panel a. of Figure 3.A.1, which plots the age-size coefficients for the exit regression representing job destruction rates for each age-size class.[37] Exit rates are especially high in the first full year of existence, from age 1 to age 2, and then rapidly decrease with age. Five years after entry, exit rates have approximately halved, but they are still significantly higher than for incumbents, i.e. firms older than six years. The ordering of the lines for different size classes further shows that exit rates decline with size within every age cohort. The same pattern holds for each age group and is even true for incumbents. These results suggest that the selection process of the passive learning model—which predicts market exit of the least efficient and therefore the smallest firms—unfolds quickly in the first years after entry.

Panel b. of Figure 3.A.1 shows that a second prediction of the passive learning model is also borne out in the Belgian data. Surviving young firms exhibit high growth rates in the early years after entry, but growth slows down rapidly with age. In contrast with the exit probabilities which decline at a relatively constant pace, the growth slowdown is most pronounced in the first few years. The average growth rate declines convexly as it converges to a constant steady state. On average, surviving young firms at age 6 still show a positive growth rate of 4 percentage points while the average incumbent does not show any employment growth.

---

[37] Recall that all regression coefficients are estimated using employment weights.

Much higher growth rates of young firms—which are overrepresented in smaller size classes—induce a negative relationship between growth and size in a cross-section of firms of all ages. Such a relationship has often been documented in the literature and it is also what we find for Belgium, as shown by the 'all firms' line in Figure 3.1 in the text. Average growth rates among all firms surviving from year t-1 to t decline monotonically with the current size of the firm. As incumbents dominate this population, the absolute growth rates are rather low, especially beyond the first two size classes.

It is instructive, however, to show the size-growth relationship separately for young firms that entered the sample at most five years ago, and older firms. The dashed line at the bottom of Figure 3.1 shows low growth rates for incumbents regardless of firm size. For them, absolute employment growth is proportional to the current size of the firm, confirming an empirical regularity found in many previous studies. In contrast, growth rates for young firms are not only higher, they clearly increase with size.

Except for this last finding for young firms, all patterns described so far are in line with results from other empirical studies based on large-scale firm-level datasets, even when no or little attempt has been made to distinguish between what we have labelled *de novo* and spurious entrants. It suggests that most patterns are fairly robust to less accurate identification of truly new and young firms.[38] The positive relationship between growth and size that we observe among young *de novo* firms, however, is not replicated in the full sample of administrative entrants. Instead, as indicated by the light gray line in Figure 3.1, the raw, administrative data suggest that small young firms have higher growth rates than larger ones. In Section 3.5.3, we showed that the difference is even more pronounced when growth rates are estimated conditional on age, that the pattern of the solid black line is robust, and how spurious entry biases the estimated relationship.

3.E Why do many studies find a negative relationship?

Several reasons why previous studies did not find the same positive relationship between growth and size of young firms have been mentioned briefly in the text. This section provides a point by point discussion. A first reason is that not all

[38] The patterns in both panels of Figure 3.A.1 and those for incumbents and all firms in Figure 3.1 in the text are qualitatively the same when calculated using the raw administrative data, reported in Geurts and Van Biesebroeck (2014).

studies condition on age, which is crucial. For example, Dunne et al. (1989) find a negative relationship but lump all firms up to age 5 in one group. Similarly, Mata (1994) finds a significant negative relationship only when young firms up to age 4 are pooled into one age class. Given the important share of young—on average high-growth—firms in smaller size classes, while larger size classes contain almost exclusively older—low-growth—firms, composition effects induce a negative relationship if firms of different ages are pooled. Pooling across all firms, incumbents and young firms, we also found a negative relationship in our dataset, see Figure 3.1 in the text.

A second reason is the inherent negative bias induced by the conventional base year classification. Most recent studies use a base year classification but control for potential bias using various other solutions than to one presented in this paper. Hence it remains unclear to what extent the difference in results is explained by different methodologies. One solution adopted by Mata (1994) is to omit all firms that enter with fewer than 10 employees to avoid truncated growth rates of the smallest firms. The same solution adopted to our sample of *de novo* entrants would imply to exclude 98.5 percent of the firms at entry. It is questionable whether the growth patterns of the 1.5 percent largest entrants are representative for those of total population of new firms entering the market. Evans (1987a) and Lotti et al. (2003) do include entrants of all sizes and use other estimation techniques to control for sample selection bias. Still, they report an inverse growth-size relationship for young firms even given age. Importantly, however, the also find convergence towards proportional growth rates for older firms, as we do.

A third reason is that spurious entrants are generally not adequately filtered out from the dataset. Since they are misclassified older firms, their growth rates tend to be much lower, resembling those of incumbents.[39] As spurious entrants dominate in larger size classes, they introduce a downward bias in post-entry growth rates that is strongly increasing with firm size. This effect is shown in Figure 3.A.2 which replicates Figure 3.3 on the full sample of administrative entrants, and on the subsample of spurious entrants that we filtered out. Panel b. shows the incumbent-like growth patterns for spurious entrants. They only grow faster than incumbents in the first recorded year and the positive growth-size relationship is not present for any age cohort, in line with the evidence for incumbents. The results confirms that the administratively recorded age of these firms is unrelated to actual firm age. Spurious entrants are already in a more

[39] In Geurts and Van Biesebroeck (2014) we show growth rates in all age-size classes separately for spurious entrants which highlights their uniformly low growth rates.

advanced stage of the selection process with less need for size adjustments. Panel a. shows growth rate estimates based on the raw sample of administrative entrants, before spurious entrants are filtered out. The downward bias spurious entrants introduce in growth rates of young firms is hardly noticeable in the smallest size classes where we showed that the share of spurious entrants is negligible. Yet, in larger size classes where spurious entrants represent the majority of administratively recorded entrants, their low growth rates swamp the high growth rates typically observed for *de novo* entrants. It obscures the positive relationship between growth and current size and even reverses it at age 2. Growth rates seem to be size invariant already from age 3 onwards.

Misclassified exits have a similar effect on the estimated pattern. Larger entrants that grow strongly are more likely to be involved in a restructuring that changes their firm ID, but is not economic exit. Some firms reorganize to cope with higher than expected growth rates, for example by splitting off some activities or adopting a different administrative structure. One incentive to split-up activities into smaller units, for example, is to remain below the size threshold of 100 employees above which firms are submitted to more stringent legal obligations.[40] Other firms are taken over by rivals that see the growth potential. Misclassifying such events involving large firms that grow strongly as exits obscures the positive size-growth relationship.

Finally, most previous studies, e.g. Evans (1987a), Dunne et al. (1989), Mata (1994), Lotti et al. (2003), focus on the manufacturing sector where the positive relationship is weaker also in our dataset. We find the increasing relationship to be most pronounced in the sectors of 'business support services' and in 'mixed business and household services,' where entry costs are often lower—graphs by industry are shown in Geurts and Van Biesebroeck (2014). Firms in these sectors can easily enter with a very small size and gradually adjust to optimal scale. Consistent with a higher minimum efficient scale in manufacturing, we find firms to enter with higher average size and show a much weaker size-growth relationship.

---

[40] 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).

Chapter 4

Employment performance following takeovers

Abstract

The merger literature has documented many motivations for takeover activity, which may lead to either employment expansion or contraction of the merged company. In public perception, however, mergers and acquisitions are associated with dramatic job loss. Empirical studies based on takeovers by listed firms in Europe have confirmed this view. Are these examples of publicly announced acquisitions representative for takeover activity undertaken by a wide range of other firms?

To shed light on this question we focus on a comprehensive sample of takeovers in the Belgian domestic market. We investigate how the individual characteristics of both the acquiring and the acquired firm affect the decision to engage in a takeover and subsequent employment growth.

The results indicate that takeovers have a small negative impact on employment growth of the merged entity that is persistent in the three post-merger periods. The adverse employment effect is mainly attributed to takeovers undertaken by small acquirers. For large acquirers we find substantial variation in post-merger employment growth suggesting that workforce rationalizations are not the dominant motivation for takeover activity. In particular, we find suggestive evidence that takeovers targeted at high-growth firms have a positive impact on firm employment growth.

JEL Codes: J23, L23

Keywords: Mergers; Acquisitions; Panel data; Labor demand

This chapter is joint work with Johannes Van Biesebroeck. I would like to thank Eric Bartelsman, John Earle, Joep Konings, Otto Toivanen for valuable comments and suggestions for improvement, as well as participants of the PhD seminar at the Center for Economic Studies KU Leuven.

4.1 Introduction

Examples of mass layoffs following takeovers of large companies have led to the popular perception that mergers and acquisitions lead to substantial workforce reductions. Even if only a small fraction of firms engage in takeover activity, our dataset shows that each year, more than 6 percent of employees in the Belgian private sector are working in a company that is involved in a takeover. If takeovers do significantly reduce the firm's demand for labor, the consequences for aggregate employment may be considerable.

This paper investigates the impact of takeovers on employment in the merged entity using a comprehensive sample of 2259 domestic takeovers in the Belgian private sector in 2007-2012. We investigate how the individual characteristics of both the acquiring and the acquired firm affect the decision to engage in a takeover and subsequent employment growth. The results indicate small but significant adverse effects on employment growth which persist for several years after the merger. We show, however, that the negative effect is not universal and strongly depends on the size, previous growth and industry characteristics of the acquirer and the target firm.

Predictions from merger theory and empirical evidence of the effects on firm employment are ambiguous. A purely anti-competitive merger reduces output and thus employment. By eliminating competition between the two companies, the integrated firm may exploit its market power and substantially increase the price of its product at the same time reducing output. A merger or takeover that increases labor productivity without changing the level of output, will reduce employment as well. Theory provides ample reason to assume that the vertical integration of firms leads to gains in labor productivity (Lafontaine and Slade, 2007). Productivity increases can for instance be realized by production cost savings, arising from economies of scale or from efficiently reallocating production and workers across the integrated firm. Conyon, Girma, Thompson and Wright (2001, 2002) and Gugler and Yurtoglu (2004) indeed find evidence that acquisitions undertaken by listed UK and European firms respectively lead to increased labor productivity and substantial job loss in the merged company.

The merger literature, however, has described many other motives that drive takeover activity (Jensen 1988). Takeovers aiming at synergy gains, the improvement of management and control, or tax incentives may as well lead to employment losses as to the growth of the firm both in output and in jobs. Mergers may also occur in response to non-profit maximizing motives (Jensen 1986; Roll

1986). Evidence of a positive employment impact has been found by Brown and Medoff (1988) for domestic takeovers in the state of Michigan. Margolis (2006) presents evidence for France that although worker displacement increases in the short term, it is significantly lower than in non-merging firms three years after the takeover. Gugler and Yurtoglu (2004) showed that takeovers undertaken by listed firms in the U.S. have no adverse effect on employment of the merged entity.

This paper contributes to previous literature in three ways. First, takeovers are defined as the integration of two independent employer firms into a single legal unit. This setting enables us to explicitly concentrate on the employment effects of merging separate workforces into a larger combined entity. Our approach differs from a related stream of literature that analyses the impact of ownership changes on the performance of the target firm only. In these studies, the impact on employment of the acquiring firm remains out of scope and job reallocation across the merged entity after the transaction is regarded as job gains or job losses. Our sample also contrasts with Conyon et al. (2002) and Gugler and Yurtoglu (2004), who focus on equity transactions between firms and include takeovers by listed companies only. Even if the two firms are brought under common ownership, they may continue as separate legal entities after the transaction. Second, we provide a comprehensive picture of takeover activity by including mergers between both small and large firms. This enables us to investigate how the employment outcomes depend on the size of the acquirer and on the relative size of the target in the merged entity. Third, and most importantly, we take into account that the characteristics of both the target and the acquirer, and the specific combination between the two affect the decision to engage in a takeover and subsequent employment growth. As a counterfactual for the takeovers, we use pairs of firms which match on the combined pre-merger characteristics of the target and the acquirer but continue as independent firms. This setting has two advantages. First, in previous studies, non-merging firms with similar characteristics as the larger merged entity have been used as a control group. If we aim at comparing the potential employment outcome of the merged firms in the absence of a takeover, pairs of non-merging firms provide a more valid counterfactual. Second, our approach enables us to assess how the combined pre-merger characteristics of the acquirer and the target lead to differential employment outcomes. In the present version of the paper, we make a first attempt to link these pre-merger features to different motivations for takeover activity.

We rely on estimation techniques from the treatment effects literature to control for the fact that firms do not randomly select into a takeover. This enables us to construct a counterfactual pair of firms with similar characteristics as the acquirer and the target up to three years prior to the takeover. In particular, we consider such features as pre-merger size, previous growth, industry, and the corporate structure of the two firms. We also make use of a set of record linking methods to control for additional firm restructurings and control changes that may occur in the pre- and post-merger periods. The potential bias this creates in firm employment histories has been largely neglected in previous studies. We show that firms in a takeover are seven times more likely to be involved in an additional restructuring than other firms, and that ignoring these events leads to a substantial underestimation of post-merger growth performance. Finally, in line with Brown and Medoff (1988), we estimate the direct effect of takeovers on employment in the merged entity. Conyon et al. (2002) have estimated the firm's derived labor demand conditional on output and wages, and distinguished between changes in labor efficiency and direct employment effects. Our sample does not include information that allow for a similar analysis.

We find that takeovers, on average, lead to significant but small reductions in employment growth of the merged entity. In the year of the transaction, employment growth is 2.4 percentage points lower than it would have been in the absence of a merger. This adverse effect persists for a substantial period of time. Growth reductions continue to be larger than 2 percentage points in the three years after the transaction. The negative employment impact is most obvious for takeovers undertaken by small acquirers in all industries. In other subsamples, however, we find suggestive evidence that workforce rationalizations are not the dominant motivation for takeover activity. This is most pronounced for takeovers undertaken by large acquirers and takeovers of high-growth targets. Takeovers by large companies lead to smaller decreases in employment and exhibit substantial variation in post-merger employment patterns. Acquisitions of high-growth firms have a more positive impact on employment growth than other takeovers, which is in line with the literature suggesting substantial synergy gains from this type of takeovers. Finally, we do not find takeovers of firms in related and unrelated industries to lead to differential employment outcomes. This result reflects recent evidence by Atalay, Hortaçsu and Syverson (2014), which suggest that vertical and horizontal integration do not fundamentally differ.

This paper is organized as follows. Section 4.2 provides an overview of previous empirical studies. Section 4.3 describes our empirical models and estimation methods. Section 4.4 discusses the data and provides summary statistics. Section 4.5 presents the results and section 4.6 concludes.

4.2 Literature

A large number of studies have investigated the impact of mergers and acquisitions on employment of the acquired plant or firm, and are based on samples of various types of control changes including plain ownership changes (Bhagat et al. 1990; Lichtenberg and Siegel 1991; McGuckin and Nguyen 2001). Some studies have estimated the impact of a foreign takeovers in particular, i.e. whether employment growth of the domestic plant is positively or negatively affected when it is acquired by a foreign company. Overall, the results indicate that takeovers have no to small negative effects on employment of the acquired plant, and that the impact varies greatly across sectors. Girma and Görg (2003), for example, show that foreign takeovers in the UK electronics industry reduced employment growth in the domestic plants, but that there was no significant effect for the food sector. Lehto and Böckerman (2008) show for Finland that foreign acquisitions had an adverse effect on employment in manufacturing plants, but not in services. For plants that are taken over by a domestic company, by contrast, the study finds a consistent negative employment effect.

Looking at employment changes in the acquired plant or firm is, however, only half the picture. Takeovers may also affect employment of the acquiring firm, and jobs may be relocated across different plants of the merged company. Studies that focus on the target firm only consider job reallocation that occurs between the target and the acquiring firm as employment gain or loss, and disregard employment changes at the level of the acquiring firm. A more consistent approach, therefore, is to consider employment changes in both the target and the acquirer. Brown and Medoff (1988), Conyon et al. (2002) and Gugler and Yurtoglu (2004) have adopted this by estimating the employment impact at the level of the combined entity, while Margolis (2006) observes employment in the two separate entities both before and after the merger.

Conyon et al. (2002) and Gugler and Yurtoglu (2004) focus on acquisitions by listed companies and find that takeovers lead to either no employment changes or significant job losses in the years following the transaction. Reductions in the firm's workforce are generally attributed to takeovers motivated by cost savings. The transaction provides an opportunity for organizational restructuring, the more efficient use of labor, and adjustment to the new optimal employment level of the merged firm. Conyon et al. (2002) find evidence for increased labor efficiency going with significant rationalizations in the use of labor in a sample of mergers and acquisitions of 277 listed firms in the UK in 1975-1996. They show negative employment effects in the order of -8 to -19 percent in the takeover period, which are particularly pronounced for two types of control changes: in the case of related mergers, where labor efficiency gains from increasing returns to scale are likely to be more substantial than in unrelated mergers (Dutz 1989); and in the case of hostile mergers, where new managers with no ties with current employees are less reluctant to renegotiate existing labor contracts (Shleifer and Summers 1988). The study, however, neglects employment in units that are divested after the merger, and is therefore likely to overestimate job losses attributed to the takeover. Gugler and Yurtoglu (2004) do control for divestiture activity in their study of 646 large mergers and acquisitions in the U.S. and Europe in 1987-1998. They argue that if mergers indeed provide an opportunity to adjust to a firm's optimal employment level, workforce reductions following takeovers are more likely to occur when firms carry excess labor. Job losses due to mergers should then be more pronounced in rigid labor markets where high labor adjustment costs prevent the pre-merger entities from operating at their optimal employment level. Holmes and Schmitz (2010) have formulated this argument in more general terms as an opportunity cost of adopting new management practices. In line with this hypothesis, Gugler and Yurtoglu (2004) find that mergers and acquisitions in European countries lead to significant reductions in labor demand of the order of -10 percent compared to the pre-merger level, while they find no adverse effects on company employment in the U.S.

Labor cost savings are, however, only one of the strategies to realize gains from takeovers. A large number of other potential benefits that drive takeover activity have been described in the merger literature, including such factors as increased market power, synergy gains, economies of scale, improved managerial competence, tax benefits, and deregulation (Jarell et al. 1988; Jensen 1988). Mergers may also occur in response to non-profit maximizing motives. Several theories explain why decision makers may have an incentive to acquire another firm even if it reduces the firms' profits. The free cash flow theory describes why managers may choose to expand the firm beyond its optimal size (Jensen 1986), and the hubris theory argues that targets of realized takeovers are, on average, overvalued (Roll 1986). The wide variety of motivations that explain takeover activity imply that the consequences for employment growth are highly ambiguous. One example of a study that has found a positive effect on employment of the merged entity is Brown and Medoff (1988), which uses a sample of 438 takeovers including also smaller firms in 1978-1984 in the state of Michigan. Their results suggest that mergers lead to significant increases in employment in the post-merger periods, especially when the comparison is not restricted to surviving firms only but also accounts for job losses due to exits.

4.3 Empirical model

4.3.1 Basic models

We estimate the impact of mergers and takeovers on employment of the merged entity in the year of the transaction and the next three years. In line with the aforementioned empirical studies, the employment series for the takeovers used in the estimations apply to the employment level of the combined entity, both before and after the transaction. Divestitures and additional changes in the firm structure are accounted for by an employment imputation procedure discussed in Section 4.4.1. We note that the distinction between a merger and a takeover is essentially a legal one without a clear-cut difference in an economic sense. Similar to previous work, we do not discern between these two types and use the terms interchangeably.

Comparing employment growth of firms that have merged and firms that have not would yield invalid estimates of the causal effect of takeover activity on firm employment. The reason is that firms do not randomly select into a takeover. Firm size, the level of competition in the industry, and other characteristics affect the decision to engage in a takeover. Moreover, employment growth and takeover activity are unlikely to be independent. Previous success, for example, is likely affect both the takeover decision and subsequent growth performance. Conyon et al. (2002) have proposed to solve this endogeneity problem by relying on an instrumental variables approach. As they are interested in estimating the simultaneous changes in firm employment, output and wages in subsequent periods after the takeover, they use a generalized method of moment's estimator, which exploits the dynamic panel features of the sample. In this paper, we opt for another estimation strategy using the counterfactual framework of the treatment effects literature. A discussion of this type of estimators is provided by Imbens (2004) and Imbens and Wooldridge (2009).

The goal of treatment effects estimators is to utilize covariates to make 'treatment' and outcome independent once we condition on those covariates. In our setting, treatment equals the takeover event. The causal effect of a takeover on employment can then be reformulated as the comparison between employment growth of firms involved in a takeover and the potential outcome if the firms would not have been merged. The goal of the estimation approach is to construct a valid counterfactual for the treated firms by selection-on-observables, i.e. the identification of observable covariates that are related to takeover activity and employment growth. If the covariates are well-specified such that the selection into takeover is random after conditioning on these covariates, treatment effects estimators yield unbiased results.

Before turning to the details of the regression, we discuss the estimated treatment and outcome models in their basic forms. In previous empirical work, the impact of takeovers on employment is modelled by regressing the logarithm of employment on a dummy variable for takeover and a set of covariates. We follow this approach with one modification. The choice of the dependent variable technically restricts the analysis to surviving firms only since the logarithm of the zero employment level of firms that exit is not defined. As our sample includes a considerable number of small firms, we want to allow for the possibility of exit and associated employment loss in the post-merger years. Exit is inversely related to firm size and job destruction due to exit among small firms has been found to be considerable (Davis, Haltiwanger and Schuh 1996a). If exit is related to takeover activity and firms that exit are excluded from the estimation, the impact of takeovers on employment will be biased. To include the possibility of exit, our dependent variable is defined as the firm-level employment growth rate. Following Davis, Haltiwanger and Schuh (1996b), growth rates are calculated as employment changes relative to the average of employment at the beginning and end of the period considered. Denoting employment of firm *i* in year *t* as E_{it}, the 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.[1] Given this definition, our basic regression has the following form:

$$g_{it} = \beta M_{it} + \sum_k \gamma_k X_{it}^k + \gamma_t + \varepsilon_{it} \qquad (1)$$

The dummy variable M_{it} takes a value of one if firm *i* is involved in a takeover in period t-1 to t. X_{it}^k are a set of observable covariates and γ_t are year dummies that control for business cycle effects. β is our coefficient of interest. If the model is correctly specified, $\beta * 100$ represents the percentage point difference between

[1] This growth rate is close to the more commonly used logarithmic growth rate $g_{it} = \ln(E_{it}/E_{it-1})$, especially for values between -1 and +1. Both measures show expansion and contraction symmetrically. Symmetry is a crucial feature for estimating mean growth rates of small firms, as their employment fluctuates widely.

the mean employment growth rate of the merged firms *i* and the outcome if they would not have been involved in a takeover.[2]

As noted above, the treatment effects estimators we will use rely on the selection-on-observable assumption, which means that conditional on the set of covariates X_{it}^{k} employment growth is independent of takeover activity. This assumption implies that systematic differences in employment growth between firms involved in a takeover and other firms with the same values for the covariates are attributable to the takeover. In this framework, g^1 would denote employment growth of a firm if it is involved in a takeover, and g^0 the outcome if it is not. Using this notation, we are interested in $g^1 - g^0$ for the merged firms, i.e. the difference between their observed employment growth and their *potential* outcome if they were not involved in a takeover. Clearly, we do not observe g^0 of the merged firms. Yet the conditional independence assumption can be reformulated as $E\ (g^0|X, M = 1) = E\ (g^0|X, M = 0)$, or the expected conditional growth rate of firms involved in a takeover would have been the same as that of other firms in the absence of a takeover. Based on this equality, treatments effects estimators use the conditional outcomes of firms that are not involved in a takeover as the counterfactual outcome g^0 for firms involved in a takeover. The effect of a takeover on employment growth of the merged firms can then be defined as $E\ (g^1 - g^0|\ M = 1)$, otherwise known as the average treatment effect on the treated.

The second assumption of treatment effects estimators to be valid is the so-called common support or overlap condition. Only observations that have a positive probability of being both treated and non-treated should be included in the analysis. The estimators are biased if, conditional on the covariates, the probability of being involved in a takeover equals either zero or one. Although current sophisticated estimators partially ensure that the counterfactual observations are chosen from the region of common support, a large subset of firms in our dataset are clearly incomparable with the firms involved in a takeover. More specifically, small and young firms are rarely involved in a takeover, while they constitute the majority of firms. We therefore restrict the sample to takeovers taking place in period t-1 to t of which the acquiring firm has at least 10 employees and is at least four years old at the time before the transaction (t-1); and the acquired firm has at least 2 employees and is at least 1 year old in t-1. The subset of control firms selected for the analysis reflects these conditions. This initial

---

[2] For small values of β and conditional on initial size, β*100 equals the percentage difference between the employment levels of the treated and untreated groups.

reduction of the sample can be considered as a pre-selection on observable characteristics.

We implement three treatment effects estimators to check the robustness of the results. In line with Conyon et al. (2002) and Gugler and Yurtoglu (2004), the first estimator compares the performance of takeovers with firms that have similar pre-merger characteristics as the combined entity. We apply a propensity-score matching estimator that matches firms on the estimated probabilities of being involved in a takeover. A fixed set of observable covariates is defined at the level of the combined entity to estimate the selection into takeover and the employment outcome. Given the minimum size and age restrictions we imposed for the acquirer and target, firms that are smaller than 10 employees and younger than 4 years in a given period *t*-1 are left out as potential counterfactuals.

A drawback of the combined-entity approach is that it disregards that takeovers are combinations of two firms with different characteristics before the merger. A large pharmaceutical company taking over a small high-growth IT firm, will reflect a different merger motivation and presumably have a different employment impact than when a medium publishing company merges with a medium retail bookseller. Yet the combined-entity method treats them as similar events. Moreover, if employment growth depends on firm size and previous growth, using counterfactual firms of similar size and lagged growth as the combined entity, will bias the results.

The second estimator we apply is a first step towards the integration of individual firm characteristics in the estimations, and takes into account the features of the acquirer in the takeover decision. We use the inverse-probability-weighted regression-adjustment estimator and define a different set of covariates in the selection and the outcome model. The treatment model is estimated as a function of the characteristics of the acquirer, while the outcome model is defined as before as a function of the characteristics of the combined entity. Here as well, the predefined size and age restrictions imply that counterfactual firms smaller than 10 employees and younger than 4 years in *t*-1 are left out from the estimations.

Our third estimator fully decomposes the combined-entity approach, and models how the characteristics of both the target and the acquirer, and the specific combination between these two affect the decision to engage in a takeover and subsequent employment growth. We use stratified matching to compare the performance of takeovers with that of pairs of firms, of which one matches the pre-merger characteristics of the acquirer, and the other one the characteristics of the target. The sample is first partitioned into cells corresponding to subclasses defined by discrete covariates of the acquirer and the target. For each cell, control units are selected from the universe of pairs of firms which exactly match on these covariates. The outcome model is then estimated as a function of the characteristics of the combined entity and the cell fixed effects.

Before implementing the estimators, we investigate which firm-level observables affect the selection into a takeover by estimating the following treatment model:

$$M_{it} = \gamma_p \, g_{it-1}^{p} + \sum_k \gamma_k \, X_{it-1}^{k} + \gamma_c C_{t-1} + \gamma_s + \gamma_t + \varepsilon_{it} \tag{2}$$

where g_{it-1}^{p} is the lagged growth rate of firm *i* measured over either one or two pre-merger periods $g_{it-1}^{p} = \left(E_{it-1} - E_{it-p}\right)/\bar{E}_{it-1}$ (p={2,3}). X_{it-1}^{k} are firm-specific variables measured at time t-1, C_{t-1} is a time-variant measure of concentration at the detailed industry level, γ_s are sector fixed effects, and γ_t are time dummies. The estimated coefficients provide insights in the firm and sector specific variables that affect the probability of being involved in a takeover, and how the timing of the takeover depends on previous growth performance. Equation (2) is estimated by maximum likelihood using a logistic model.

4.3.2 Estimated equations

Based on preliminary estimates of equations (1) and (2), the set of observable covariates to be included in the final estimations was selected. The definitions of these variables are provided in Appendix Table 4.A.1. As noted before, propensity score matching uses only one set of covariates to estimate the probability of treatment and the employment outcome. The inverse-probability-weighted regression-adjustment estimator defines the treatment and the outcome model with different sets of covariates. The stratified matching approach first uses a set of discrete covariates to match takeovers with counterfactual pairs of firms, and estimates the outcome model using additional covariates.

Propensity-score matching

Propensity-score matching chooses counterfactual firms for the takeovers based on the estimated predicted probabilities of being involved in a takeover, the so-called propensity score. This measure indicates the degree of similarity between two firms based on observable characteristics that affect takeover activity. The propensity-score estimator we implement is based on the following treatment model:

$$M_{it} = \alpha E_{it-1} + \gamma_3 g_{it-1}^3 + \gamma_g G_i + \gamma_f F_i + \gamma_c C_{t-1} + \gamma_s + \gamma_t + \varepsilon_{it} \qquad (3)$$

where E_{it-1}, represents the employment level of firm i in t-1 and g_{it-1}^3 is the growth rate of firm i measured over the two previous periods $g_{it-1}^3 = (E_{it-1} - E_{it-3})/\bar{E}_{it-1}$, with $\bar{E}_{it-1} = (E_{it-1} + E_{it-3})/2$. The dummy variable G_i indicates whether the firm is part of an enterprise group and the dummy F_i whether it is engaged in FDI in the sample period. C_{t-1} is the industry concentration ratio at time t-1 expressed as the employment share of the four largest firms at the Nace 3-digit level[3], and γ_s and γ_t are sector and year fixed effects.

In the estimator we apply, we impose that each firm involved in a takeover should be matched with at least five similar firms, the so-called 'nearest neighbors'. The maximum difference in propensity scores between matched firms we allow is set at 0.05 standard deviation. Within this caliper, our actual estimations find between 5 and 20 matches. The effect of takeover activity on employment growth is then computed by taking the average of the difference between the observed outcomes for each firm involved in a takeover, and the potentials outcomes computed as the average of the outcomes of the matched firms.

Inverse-probability-weighted regression-adjustment

The inverse-probability-weighted regression-adjustment estimator fits weighted regression models of the outcome for each treatment level, where the weights are the estimated inverse propensity scores obtained by the treatment model. The regression coefficients are then used to predict two outcomes for each firm involved in a takeover, one for each treatment level. The difference between the averages of these outcomes estimate the effect of takeover activity on employment growth of the merged firms. Imbens and Wooldridge (2009) show that this estimator has a double-robust property. It yields consistent estimates if either the selection model or the outcome model is correctly specified. Exploiting this flexibility, we define different sets of covariates to fit the separate models. The treatment model is estimated as a function of the characteristics of the acquirer, while the outcome model is defined as before as a function of the characteristics

---

[3] In our sample, the Nace 3-digit level corresponds to 166 separate industries.

of the combined entity. The coefficient estimates of the treatment model will therefore more closely reflect how the decision to takeover depends on characteristics of the acquirer than it is the case in treatment model (3), where the covariates refer to the combined entity.

The treatment model that estimates the propensity scores takes the following form:

$$M_{it} = \alpha e_{it-1} + \alpha^2 e_{it-1}^2 + \gamma_{pa} PA_i + \gamma_f F_i + \gamma_c C_{t-1} + \gamma_s + \gamma_t + \varepsilon_{it} \qquad (4)$$

Where e_{it-1} represents the logarithm of employment of firm *i* at time *t*-1 and allows the selection into takeover to be dependent on firm size measured at the time before the transaction. Two covariates indicate how the decision to takeover is related to other control activities of the firm. The dummy variable PA_i indicates whether firm *i* is a parent firm, i.e. owns at least 50 percent of the shares of another Belgian firm in the sample period, and the dummy F_i whether it is engaged in FDI. As noted above, the covariates for the takeovers represent the values for the acquiring firm only. The other variables are defined as above.

The outcome model has the following form:

$$g_{it} = \alpha E_{it-1} + \gamma_3 g_{it-1}^3 + \gamma_g G_i + \gamma_f F_i + \gamma_c C_{t-1} + \gamma_s + \gamma_t + \varepsilon_{it} \qquad (5)$$

where the variables are defined as above. The control variable E_{it-1}, representing the employment level of firm *i* in *t*-1, allows growth rates to be dependent of firm size and to deviate from Gibrat's law of proportionate growth (Sutton 1997). As noted, equation (5) is estimated as weighted regressions for each treatment level, i.e. for $g_{it} = \{g_{it}^0, g_{it}^1\}$.

Stratified matching

Our third estimator strongly departs from the combined entity approach applied in previous papers by taking into account that the characteristics of both the target and the acquirer, and the specific combination between these two may affect the decision to engage in a takeover and subsequent employment growth. In particular, we consider such features as pre-merger size, previous growth, detailed industry, and the corporate structure of each of the two firms. As a counterfactual for the takeovers, we use pairs of firms with the same combined pre-merger characteristics. Our approach further differs from the two other treatment effects estimators that use the propensity score of the treated firms to select a counterfactual. The propensity score is a linear measure indicating the degree of similarity between two firms based on a combination of observable covariates. Our counterfactual pairs of firms, by contrast, will be chosen such that they exactly match the takeovers with respect to all chosen covariates. Exact matching eliminates all imbalances (i.e., differences between the treated and control groups) beyond the level defined by the combined covariates. The remaining differences are thus all within small cells of firms with the same values for the covariates. In terms of exogeneity, our estimator relies on the assumption that employment growth is independent of takeover activity conditional on the cell characteristics, which are defined by the full interaction of the individual and the combined characteristics of the acquiring and target firm. As the choice of counterfactual firms is based on more precisely defined similarities with the takeovers than in the two previous estimators, our stratified matching approach will yield more reliable and more precise estimates of the takeover effect.[4]

Our estimation procedure proceeds in two steps. First, the sample of takeovers and the universe of pairs of potential control firms is partitioned into cells defined by discrete covariates of the acquirer and the target. Next, the outcome model is estimated as a function of the characteristics of the combined entity and the cell fixed effects.

The sample of paired firms includes all combinations of two firms of which one fulfills the predefined characteristics of an acquirer (at least 10 employees and at least four years old in t-1) and the other those of a target (at least 2 employees and at least 1 year old in t-1). The combined sample of takeovers and pairs is subdivided into mutually exclusive cells (or 'strata') c_l, using 7 discrete stratification variables $x_1, \dots, x_7$. The choice of these variables is based on the results of our previous analysis, reflecting the characteristics of the acquirer and the target that may affect the decision to engage in a takeover and subsequent growth. The stratification variables are defined as follows:

$x_1 = \{2007, \dots, 2012\}$ is the year at the time before the takeover (t-1),

$x_2 = \{011, \dots, 960\}$ is the Nace 3-digit industry code of the acquirer (166 industries),

[4] Our stratified matching approach is closely related to various exact matching estimators described in the literature. It corresponds to nearest neighbor matching when it is enforced that all covariates are discrete and match exactly (Imbens 2004). With exact matching on discrete covariates, the nearest neighbor matching estimator reduces to an average of differences in cell means. The regression adjustment estimator with fully interacted discrete covariates reduces to the same average of difference in cell means. Our strategy is further related to stratified matching as described by Anderson, Kish and Cornell (1980). Exact matching estimators have the property to reduce bias more and gain in precision as the number of subclasses is increased.

$x_3 = \{011, \ldots, 960\}$ is the Nace 3-digit industry code of the target (166 industries),

$x_4 = \{1, 2, 3, 4\}$ is the size quartile of the acquirer defined within each year and industry cell $(x_1 * x_2)$,

$x_5 = \{0, 1\}$ is a dummy indicating whether the acquirer is part of an enterprise group in t-2 or t-1,

$x_6 = \{1, 2, 3, 4,5\}$ is the lagged growth quartile of the target defined within each year and industry cell $(x_1 * x_3)$. Lagged growth is defined as the growth rate measured over the three periods before t-1, $g_i = (E_{it-1} - E_{it-4})/\bar{E}_{it-1}$, with $\bar{E}_{it-1} = (E_{it-1} + E_{it-4})/2$. The fifth subset includes targets that entered in one of these periods,

$x_7 = \{0, 1\}$ is a dummy indicating whether the employment share of the target in the combined entity in t-1 is larger than 25 percent.

The combination of the 7 stratification variables generates 13 226 880 possible cells c_l. Observations in cells that contain at least one takeover and one control pair of firms are retained; observations in the remaining cells are removed from the sample.[5] This leaves us with 1822 cells with at least one takeover and one counterfactual pair.[6] Of all takeovers, 2 percent for which no counterfactual pair exists drop out.

The outcome model is defined as a function of the characteristics of the combined entity and the cell fixed effects. It has the following form:

$$g_{it} = \beta M_{it} + \alpha E_{it-1} + \gamma_3 g_{it-1}^3 + \gamma_c + \varepsilon_{it} \tag{6}$$

where γ_c represents the cell fixed effect and the other variables are defined as above. The model is estimated using a weighted least squares regression, where the weights are equal to the number of takeovers in each cell.[7] The coefficient β is

[5] To restrict the number of counterfactual pairs per cell to a manageable number, we randomly select 150 acquirers within each subclass defined by x1, x2 and x4, and 150 targets within each subclass defined by x1, x3 and x6. This only reduces the number of potential counterfactuals in industries with many small firms such as construction, retail or restaurants.

[6] The minimum number of takeovers and counterfactual pairs per cell in the final sample is 1. The maximum number is 5 takeovers and 21 136 counterfactuals.

[7] This weighting scheme is not fully appropriate for estimating a model that besides the cell fixed effects also includes the continuous covariates initial size and previous growth. We will look for a more appropriate weighting scheme to be applied in the final version of the paper.

an estimate of the effect of takeovers on employment growth corresponding to the average of the effect within the cells. Note that the cells c_l correspond to the full interaction term between the 7 covariates $x_1, \dots, x_7$. The cell fixed effects are thus perfectly collinear with the covariates for time, industry, industry concentration and enterprise group used in the outcome models of the two previous estimators.[8]

4.3.3 Extensions

In time

The equations discussed above estimate the impact of takeover activity on firm employment immediately after the transaction. As labor adjustments can be slow or takeover activity can have a persistent impact on employment growth, we extend the regressions to the 3-year period after the takeover. The post-merger impact on employment will be estimated both as year-by-year changes, and as cumulated changes over the entire post-merger periods. The first provide information on the dynamic adjustment paths conditional on surviving, while the latter give insight in the long-term employment gains or losses following a takeover.

The regressions estimating employment growth over n-year periods use growth rates that take employment at time t-1 as the base year and are calculated as $g^{+}_{it+n} = (E_{it+n} - E_{it-1})/\bar{E}_{it+n}$, with $\bar{E}_{it+n} = (E_{it+n} + E_{it-1})/2$ and $n = \{1,2,3\}$. The other regression variables do not differ from the ones specified in equation (3) to (5).

The regressions estimating year-by-year employment changes use growth rates calculated over the preceding period where $g_{it+n} = (E_{it+n} - E_{it+n-1})/\bar{E}_{it+n}$, $\bar{E}_{it+n} = (E_{it+n} + E_{it+n-1})/2$ and $n = \{1,2,3\}$. Here as well, the regression variables do not change except for E_{it-1} in equation (5), which is replaced by E_{it+n-1} and controls for the size of the firm at the start of the period considered.

By subset

The effect of takeover activity on firm employment growth may fundamentally differ by the type of the transaction, the characteristics of the acquiring and the acquired firm, and by sector. Two major distinctions we focus on is the differential impact of takeover activity by small and large acquirers, and the impact by broad

[8] In the present version of this paper, the FDI variable has not yet been included in the stratified matching model.

sector groups. To investigate effect along these dimensions, we estimate separate regressions by subset.

Finally, we use a minimum distance estimator to investigate whether the impact on employment differs by a set of other characteristics of the target and the acquirer. We again include the distinction between small and large acquirers, and additionally distinguish between small and large target shares, high-growth versus other targets, and related versus unrelated mergers. The four combined characteristics lead to 16 subsets of takeovers. The estimation strategy follows Wooldridge (2010) and proceeds in two steps. First, the takeover effect on employment is estimated for each of the 16 types of takeovers separately using inverse-probability-weighted regression-adjustment. This provides us with 16 coefficients of the takeover effect. Next, the minimum distance estimator is obtained from an OLS regression of the 16 coefficients on the dummies of the characteristics. More specifically, we use a weighted least squares regression, where the weights are the squared standard errors of the coefficients obtained in the first step.

4.4 Data

4.4.1 Sample composition

We use a sample of 2259 domestic takeovers in 2007 to 2012 which are defined as the integration of two previously independent Belgian employer firms into a single legal unit.[9] This setting allows us to explicitly concentrate on the employment effects of merging separate workforces into a larger combined entity. The sample consists of all mergers that are observed in the specific period, including transactions between both small and large firms. The only size restriction we impose is that the acquirer has at least 10 employees and the target at least 2 employees at the time before the transaction.

The identification of takeovers is based on two sources. The first is based on employee-flow linkages between firms using a linked employer-employee dataset (Geurts 2016). A takeover is defined as an event where an incumbent absorbs the

[9] Independency is based on the official firm identification number. Each firm ID number is regarded as a separate firm by the Belgian law. Before the merger, each firm pays its own social security contributions, corporate taxes, and fills out individual annual accounts. After the merger, these obligations are fulfilled by the joint entity.

entire workforce of another firm and the latter is dissolved after the transaction. Similarly, mergers are identified as two firms that are dissolved and merge their workforces into a newly created firm.[10] The second source is a dataset compiled by Statistics Belgium based on official mergers and acquisition approved by the Commercial Court. It includes share deals between companies where the buyer becomes the owner of legal entity, and acquires the target's shares and assets as well as all existing liabilities and debts.[11] In line with our definition above, we include only events when two firms merge into a single legal entity. The vast majority of observations in our sample are takeovers, i.e. targets that are absorbed by a continuing incumbent. Plain mergers, where both firms are dissolved and a new merged company is created after the transaction, represent a small share of our sample.[12] In line with previous studies, we do not distinguish between 'mergers' and 'takeovers' and use the terms interchangeably. When several firms are taken over in the same year, the transactions are collapsed into one event. Takeovers in different years, however, are included as separate observations.

Employment information is based on the register of Belgian employers maintained by the National Social Security Office (NSSO). It covers all Belgian private firms with at least one employee. This dataset is also used to construct the control group of firms. The employment files cover the period from 2003 to 2012. This leaves us with an unbalanced panel including employment histories of unequal length for the takeovers in 2007-2012. Information about the control structure of firms is included using a dataset provided by Statistics Belgium.

The sample of 2259 takeovers is pre-selected on a set of characteristics to avoid measurement error and confounding sets of firms with incomparable growth patterns. Temporary agencies, which exhibit continuous reshuffling of legal entities within enterprise groups, and firms in highly subsidized sectors, where employment growth strongly depends on policy measures, are excluded

[10] More specifically, a takeover corresponds to an event where at least 50 percent of the individual employees of the dissolved firm is transferred to the incumbent. Similarly, a merger corresponds to the transfer of at least 50 percent of the individual employees of two dissolved firms into a new legal entity. Most takeovers and mergers in the sample correspond to transfers of close to 100 percent of the workforces into the combined entity.

[11] Not all takeovers are subject to the Commercial Court procedure. Asset deals, 'noiseless' mergers between firms of the same corporation, and other buy and sell operations can be executed without approval by the Commercial Court.

[12] Plain mergers represent less than 3 percent of the total sample and are not investigated separately. In the case of plain mergers, we consider the largest predecessor as the acquirer and the other predecessor(s) as the acquired firm(s).

from the analysis. We also exclude some exceptional observations where the acquiring firm is very young or very small. Young and small firms are known to have extremely high growth and exit rates compared to other firms, which would lead to outliers in the estimations (Geurts and Van Biesebroeck 2014). Finally, we exclude takeovers where the employment share of the target in the combined entity is smaller than 1 percent. These takeovers do not correspond to substantial workforce integrations, which is the focus of our study.

Technically, our sample of 2259 takeovers includes all mergers between two or more firms that occurred in period t-1 to t where t = {2007,…, 2012} and satisfy the following conditions. The acquirer has at least 10 employees and is at least 4 years old at the time before the transaction (t-1); the target is at least 1 year old and has at least 2 employees in t-1 and represents at least 1 percent of employment of the combined entity; the merged entity survives in t.

From all Belgian firms *not* involved in a takeover in 2007 to 2012, a sample of 600 000 observations is selected as potential counterfactuals for the target. The sample is predefined on the same observable characteristics: the firm has at least 2 employees in year t-1 and is in existence at least since from year t-2 to t. A firm may be included more than once in the sample if it satisfies these conditions for different values of t. A subset of 140 000 observations that meet the conditions of an acquirer is used as the potential counterfactual group for the acquiring firms. The propensity score matching and inverse-probability-weighted regression-adjustment estimators use only counterfactuals for the combined entity. The subset that meets these conditions consists of 128 000 observations. The final sample is an unbalanced panel as it does not include employment information before 2003 and after 2012.

We take great care to control for other changes in the firm structure that may occur before or after the takeover. Takeovers can be accompanied by divestitures of parts of the firm in the year of the transaction and in the post-merger periods. Moreover, firms may engage in another merger, split up, or disappear from the dataset because they change identification number. Such events are known to be a major source of bias in the measurement of firm-level employment growth based on micro-level data (Haltiwanger et al. 2013; Geurts 2016). It has been neglected in previous studies that these biases are likely to be exacerbated in the population of takeovers. We find that firms involved in a takeover have a much higher probability to be involved in an additional restructuring than other firms. Appendix Table 4.A.2 shows that 34 percent of the takeovers in our sample are involved in another event in the three years following the merger, as opposed to only 5 percent of the firms in the control sample. Previous studies, which did not control for such events, are likely to report employment growth figures which are artificially inflated or deflated.[13]

We address this methodological shortcoming in the following way. A set of advanced record linking methods is used to control for ID changes and changes in the firm structure that occur before or after the takeover period t-1 to t. The methods are based on supplementary data sources, probabilistic matching and employee-flow record linking as described in Geurts (2016). The linkage methods allow us to reconstruct the employment histories of firms involved in a takeover three years before and three years after the transaction. Our approach is to impute employment growth at the firm level by assuming the same growth rate for each firm involved in the event.[14] This imputation method treats split-offs and absorptions symmetrically and preserves the firm size distribution in the sample. It allows a more accurate estimate of the net effect of takeovers on firm employment growth. Imputation is only performed until a second event occurs. Beyond that, firm observations are excluded from the analysis since correcting for multiple events often involves a complex set of interlinked firms and imputation becomes unreliable.[15]

---

[13] Gugler and Yurtoglu (2004) partly address this problem by introducing a dummy for divestiture activity.

[14] More specifically, the approach we adopt is to first construct an aggregate event level including all firms interlinked in a given period from t to t+1. Firm-level employment in t+n with n = {1,…} is then imputed by assuming the same growth rate for each firm involved in the event. The imputation is carried out for the three post-merger periods. Similarly, backward employment imputation is applied in a given period from t to t-1 for the three pre-merger periods. More detail on the methods for identifying events and imputing employment histories is provided in Geurts (2016).

[15] In the present draft of this paper, employment imputation is only carried out for firms involved in a takeover. Firms in the control sample, were events are exceptional, are simply excluded from the analysis in the periods they are involved in an event. We hereby avoid artificially inflated or deflated growth rates for this subset as well.

4.4.2 Summary statistics

Table 4.1 shows the composition of the sample of takeovers by subsets. The middle three columns report the average size of the acquirer, the target and the combined entity at the time before the merger (t-1).[16] The last column shows the percentage share of the target in the combined entity. The first row shows that, on average, acquirers employ 191 employees at the time before the merger. They are almost 5 times larger than the firms they take over. The joint entity has an average size of 231 employees. This combined size will be the starting point for the estimations of post-merger employment growth. The rest of the table shows these summary statistics by subsets of takeovers. The definitions of the subsets are provided in Appendix Table 4.A.1.

The distinction between small and large acquirers indicates why it is interesting not to focus on takeovers by large firms only. First, smaller firms, here defined as firms with less than 100 employees, do exhibit considerable merger activity and constitute 70 percent of the sample. We select a cut-off value of 100 employees because this is a threshold for more stringent legal obligations in Belgium.[17] Second, the weight of the target firm in the transaction differs greatly between small and large acquirers. Small acquirers takeover targets that, on average, represent 30 percent of the merged entity, while in the case of large acquirers, the target share is only 15 percent. The statistics of the next subsets, by the share of the target in total employment of the combined entity, confirm this pattern. Takeovers are divided into two approximately equal by subsets by using a cutoff value of 25 percent to distinguish small and large target shares. Small target shares are associated with large acquirers and vice versa. It has been suggested that smaller acquirers show greater post-merger efficiency gains than their larger counterparts (Conyon et al. 2002). This differential impact may be related to the relative sizes of the merged firms, and we will control for both effects in the estimations.

Growth related merger theories have put forward that one motivation for acquiring another firm is to gain access to new technologies and highly specialized personnel. Moreover, high-growth firms with low liquidity and high leverage have

[16] In the case of multiple targets, the average target size is based on the sum of employment of the different targets. In the case of mergers, the largest firm in t-1 is defined as the acquirer.

[17] In Belgium, small firms do not need to file full annual accounts or install a work council (with fewer than 100 employees, turnover below 7.3 m EUR, and balance sheet total below 3.65 m EUR).

been found to be likely targets (Palepu 1986). As the merger generates substantial synergy gains for both the target and the acquirer, the employment impact can be expected to be more positive than when labor cost savings are the dominant motivation. We will examine this hypothesis by distinguishing takeovers according to the previous growth performance of the target. High-growth targets are defined as firms with an average annual growth rate of at least 8 percent in the three years before the merger. They represent 25 percent of the sample.

Table 4.1 Sample composition of takeovers in 2007-2012

	Number of obs.	Average size in *t*-1 (number of employees)			Percentage share of target
		Acquirer	Target*	Combined entity	
Total	2259	191	40	231	17.3
By size of acquirer: less versus more than 100 employees					
Small	1582	36	16	52	30.8
Large	677	554	96	650	14.8
By employment share of target: less versus more than 25% of combined entity					
Small	1171	295	27	322	8.4
Large	1088	79	54	134	40.3
By previous growth performance of target: average annual growth rate in the 3 years before the takeover less versus more than 8%					
High growth target	568	148	29	177	16.4
Other	1691	206	44	250	17.6
By type of integration: target and acquirer in same versus different Nace3-digit industry					
Related	1313	192	47	239	19.7
Unrelated	946	190	30	220	13.6
By data source					
Commercial Court	1312	217	52	269	19.3
Employee-flow method	1925	208	45	253	17.8
By type of merger					
Takeover of 1 target	2058	146	29	175	16.6
Takeover of more targets	138	919	211	1131	18.7
Merger	63	71	23	94	24.5

* In the case of multiple targets, the average size of the sum of employment of the targets is reported.

In line with previous studies, we also distinguish between 'related' and 'unrelated' mergers. Takeovers are classified into related and unrelated depending on whether the target and the acquirer belonged to the same Nace 3-digit industry.[18] Here as well, predictions from merger theory and empirical evidence on the employment consequences are ambiguous. Employment losses have been hypothesized to be more likely in related than in unrelated mergers, particularly when the industry exhibits substantial economies of scale (Dutz 1989). Atalay, Hortaçsu and Syverson (2014), however, argue that vertical and horizontal expansion do not fundamentally differ as they both aim at facilitating efficient transfers of intangible inputs. Conyon et al. (2002) and Gugler and Yurtoglu (2004) have found contrasting empirical results on this issue. Table 4.1 shows that related takeovers represent half of the sample. With respect to the average absolute and relative sizes of the firms involved, they do not differ greatly from unrelated mergers.

The bottom part of the table gives more information about the construction of the sample. The Commercial Court data file identifies 58 percent of the merger activity. Most of these officially registered mergers and acquisitions are picked up by the employee-flow method as well, which identifies 85 percent of the takeovers in the sample. Finally, the last subsets show that the majority of observations in the sample are simple takeovers of one target by one acquirer (91%). Takeovers of multiple targets in the same period (6%) and 'plain' mergers of firms that dissolve and merge into a newly created company (3%) are rather exceptional.

Table 4.2 presents further information on the sectors that will be investigated and on the control sample. In all sectors, defined by the industry of the acquirer, firms involved in a takeover are only a small fraction of active firms but they represent a considerable share of employment. The first row shows that every year, not less than 6 percent of all employees in the Belgian private sector are working in a company that is involved in a takeover. If takeovers do significantly affect the firm's use of labor, the consequences for aggregate labor demand may be considerable. Business and household services exhibit the most intense takeover activity, both in terms of the absolute number of takeovers and their share in sector employment (7.8%). Manufacturing shows the highest share of firms involved in a takeover (1.1%) and the largest average sizes of acquirers and targets. Wholesale and retail show average takeover activity, while Construction is dominated by merger activity between smaller firms.

[18] In our sample, the Nace 3-digit level corresponds to 166 separate industries.

Table 4.2 Summary statistics by sector

	Number of obs.	Share in sector population (%)	Share in sector employment (%)	Average size in t-1		
				Acquirer	Target	Combined entity
Total						
Takeovers	2259	0.5	6.1	191	40	231
Potential counterfactuals				38	12	41
Manufacturing						
Takeovers	526	1.1	7.4	263	49	311
Potential counterfactuals				56	25	59
Construction						
Takeovers	225	0.3	3.1	123	24	147
Potential counterfactuals				29	10	31
Wholesale and retail						
Takeovers	611	0.5	5.5	163	28	192
Potential counterfactuals				32	10	33
Business and household services						
Takeovers	833	0.6	7.8	185	48	233
Potential counterfactuals				41	11	44
Other (Agriculture, Accommodation)						
Takeovers	64	0.1	2.0	197	37	233
Potential counterfactuals				22	8	23

Note: Annual averages in the period of observation (2007-2012).

The last three columns also include information on the sample of other firms that are used to construct the counterfactuals for takeovers. They show the average sizes of other firms from which the potential counterfactual targets and acquirers for the stratified matching estimations are selected; and, in the last column, the average sizes of the firms that are used as potential counterfactuals for the combined entity in the propensity score matching and inverse-probability-weighted regression-adjustment estimations. Even if we pre-selected on initial size, survival, and age as described above, the average sizes of the potential counterfactuals are much smaller than those of the firms involved in actual takeovers. In each sector, actual acquirers and combined entities are more than four times larger than their potential counterfactuals, and actual targets are more than two times larger. Figure 4.A.1 in the Appendix provides a more detailed picture of the size distributions of targets, acquirers, and their potential counterfactuals. The figure shows the distributions at the time before the transaction (t-1) based on the number of employees on a logarithmic scale. Actual targets and acquirers show moderately right-skewed distributions with the highest share of firms in size classes 8 to 15 employees and 16 to 63 employees respectively. An important share of actual acquirers is located in the right tale of the distribution. Firms in the potential counterfactual groups, by contrast, show extreme right-skewed distributions with more than 75 percent of the firms concentrated in the first two size classes and almost no tale. From these statistics, one can presume that the treated and control group differ strongly with respect to other characteristics as well. A further selection-on-observables and an appropriate weighting scheme is thus necessary to obtain more reliable estimates of the effect of takeover activity on firm employment.

Figure 4.1 and 4.2 present the evolution of takeover activity by sector in the period of observation 2007-2012. In most sectors, the share of firms involved in a takeover has been relatively stable over time. Only Manufacturing, where Belgian firms suffered greatly from the decline in export during the global recession of 2008-2009, showed a temporary increase of takeover activity following the recession. Takeover activity shows much more variation in terms of employment shares (Figure 4.2), as the size of the firms involved differs strongly between takeovers. The peak in Business and household services in 2010, for example, is explained by a few very large firms engaging in a takeover in Publishing, ICT and the Financial sector.

Figure 4.1 Share of firms involved in a takeover by sector

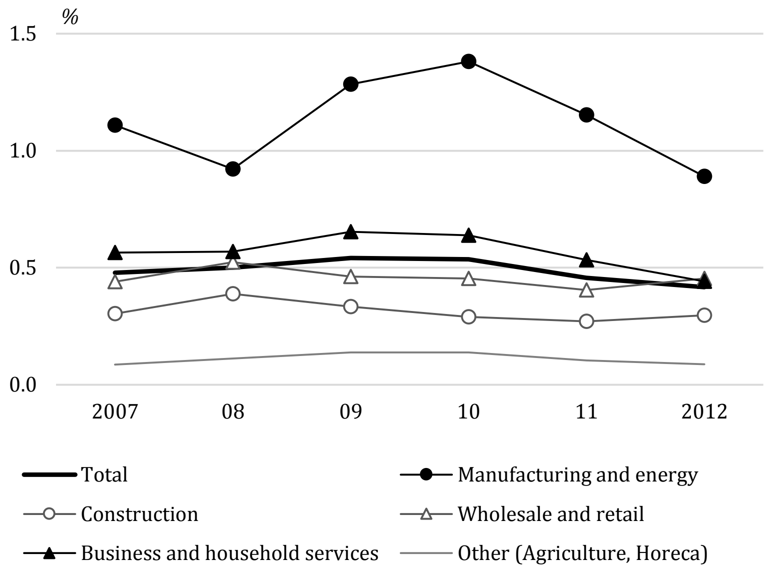
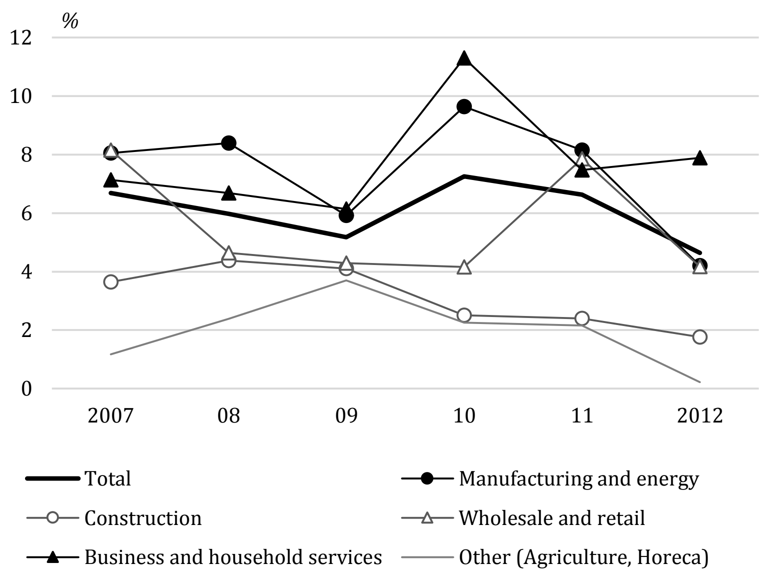


Figure 4.2 Employment share of firms involved in a takeover by sector



4.5 Results

4.5.1 Selection into takeover

To investigate which firms engage in the acquisition of another firm, we first show results for the treatment model (4) with the takeover dummy as the dependent variable. The equation is estimated by maximum likelihood using a logistic model. Empirical studies have suggested numerous firm-level variables that explain takeover activity and which are related to different hypotheses about the motivations for takeovers such as shareholder premiums, management competence, free cash flows, growth-resource imbalances, firm size, and so on (e.g. Jarell et al. 1998; Roll 1986; Palepu 1986). Other studies have found also industry level variables to be significant including industry shocks, growth, concentration, asset liquidity (Mitchell and Mulherin 1996), deregulation and antitrust relaxation (Jensen 1988).

Given the limitations of our dataset, we are able to include the following explanatory variables that have previously been found to predict takeover activity. At the firm level we include firm size and previous growth performance measured in terms of employment; and firm dummies which are informative about FDI flows and the corporate structure of the firm. Industry characteristics are picked up by fixed effects that are tested at various levels of the industry classification. Finally, we examine the significance of different measures of industry concentration at the detailed Nace 3-digit level.

The results for the treatment model are presented in Table 4.3. In the first two columns, we examine how takeover decisions depend on the size and previous growth of the acquirer. The regressions also include sector and year fixed effects. The probability of takeover increases nonlinearly with firm size. Lagged growth is positively correlated with the takeover decision, with a larger coefficient if measured over a 2-year than 1-year period. But given the large variation in pre-merger growth, the effect is not significant. The sector fixed effects (detail not provided in table) reveal that after controlling for firm size and growth, manufacturing companies are less likely to engage in a takeover than it appeared from the summary statistics in Table 4.2. Firms in Information and communication services and in Financial services exhibited the highest takeover activity. The year dummies confirm an increase in takeover activity in 2009-2010, and a decline thereafter.

Table 4.3 Which firms select into takeover
Dependent variable: dummy for takeover in period t-1 to t

	(1)	(2)	(3)	(4)	(5)
Firm-level variables					
Employment $_{i\,t-1}$ (log)	1.617***	1.617***	1.452***	1.471***	1.401***
	(0.101)	(0.101)	(0.102)	(0.102)	(0.103)
Employment^2 $_{i\,t-1}$ (log)	-0.089***	-0.089***	-0.082***	-0.077***	-0.076***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
1-year lagged growth $_i$ (t-2 to t-1)	0.012				
	(0.147)				
2-year lagged growth $_i$ (t-3 to t-1)		0.051			0.077
		(0.100)			(0.102)
Parent firm $_i$ (dummy)			1.007***		0.982***
			(0.051)		(0.051)
Enterprise group $_i$ (dummy)				0.484***	
				(0.045)	
FDI $_i$ (dummy)			0.279***		0.250***
			(0.069)		(0.070)
Industry variables					
Concentration ratio $_{t-1}$ (Nace 3-digit level)			-0.726***		-0.744***
			(0.140)		(0.162)
Herfindahl Index $_{t-1}$ (Nace 3-digit level)				-0.987**	
				(0.389)	
Sector FE	Yes	Yes	Yes	Yes	No
Nace 2-digit FE	No	No	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.097	0.097	0.117	0.102	0.123
No. of observations	130 284	130 284	130 284	130 284	130 284

Regression coefficients show results of logistic regressions. Standard errors in parentheses.

In the next two columns we add various variables on the control structure of the acquiring firm and the level of industry concentration. Column (3) shows that being a parent firm, i.e. owning at least 50 percent of the shares of another Belgian firm, significantly increases the probability of takeover. Involvement in FDI flows has positive impact as well. By contrast, a high level of industry concentration at the time before the takeover, measured as the employment share of the four largest firms at the detailed 3-digit industry level, is a negative predictor for takeover activity. Column (4) tests alternative versions of these explanatory variables. We include an enterprise group dummy, indicating whether the firm is controlling or controlled by another company, and the Herfindahl index as a measure of industry concentration. Using these alternatives does not fundamentally change the results but the pseudo-R2 indicates that the predictive power of the model decreases. We therefore use the covariates in column (3) in our final estimations.

So far, industry fixed effects have been added with dummies for eight broad sector groups. When estimating the outcome model for different subsets of takeovers, the limited number of observations does not allow us to go beyond this level. When, however, the overall impact of takeover activity is estimated, fixed effects of more detailed industries can be included to more accurately capture the specific industry characteristics that affect the takeover probability. The last column reports the coefficients of the model when the industry fixed effects are added at the Nace 2-digit level, with separate dummies for 34 industries. Compared to column (3), the sizes of the coefficient estimates slightly decrease, indicating that the detailed industry level accounts for part of the variation in the explanatory variables. All coefficients except the one of lagged growth remain, however, highly significant and the explanatory power of the model increases.

4.5.2 Effect of takeovers on firm employment growth

Tables 4.4 and 4.5 present our estimation results for the impact of takeovers on firm-level employment growth until the third year after the transaction. Table 4.4 reports the effect on year-by-year employment growth rates and Table 4.5 shows the results when growth rates are calculated over the *n*-year post-merger period using employment at time *t*-1 as the base year. The coefficient estimates in panel a. of each table correspond to the average treatment effects on the treated obtained by either propensity score matching or inverse-probability-weighted regression-adjustment. The propensity score matching estimations are based on treatment model (3), including firm-level covariates at the level of the combined entity only. Inverse-probability-weighted regression-adjustment is based on treatment model (4), which specifies control variables in terms of the characteristics of the acquirer, and on outcome model (5). The regressions for both estimators only included takeovers for which the treatment models showed that at least five matches could be found that are sufficiently similar to the takeover. We set this threshold for similarity at 0.05 standard deviation of the propensity score. Panel b. reports the coefficient estimates based on stratified matching and outcome model (6). As discussed above, this estimator uses pairs of firms as a comparison group of which both the counterfactual target and acquirer match each of the two firms involved in a takeover on a set of individual characteristics.[19]

Table 4.4 Effect of takeovers on year-by-year employment growth
Dependent variable: employment growth rate (in percent)

	Takeover period	*n*th post-merger period		
		1st	2nd	3rd
	(*t*-1 to *t*)	(*t* to *t*+1)	(*t*+1 to *t*+2)	(*t*+2 to *t*+3)
Panel a.				
Propensity score matching				
	-1.86***	-2.18***	-1.88**	0.14
	(0.32)	(0.64)	(0.81)	(0.83)
Inverse-probability-weighted regression-adjustment				
	-2.16***	-2.77***	-3.03***	-1.51**
	(0.29)	(0.59)	(0.76)	(0.72)
No. of observations	130 274	107 789	84 760	62 226
No. of takeovers	2 249	1 929	1 503	1 034
Panel b.				
Stratified matching on acquirer and target characteristics				
	-2.43***	-2.22***	-2.78***	-2.25***
	(0.25)	(0.45)	(0.54)	(0.67)
No. of observations	5 555 003	4 439 591	3 492 217	2 573 014
No. of takeovers	1 948	1 670	1 316	916

Regression coefficients show the average treatment effects on the treated. Standard errors in parentheses.

Takeovers in 2007-2012. Observations drop out from the estimations when employment is outside period of observation (after 2012). Regressions include all firms that survive or exit in a given period. Regressions do not include firms that have exited from the year after exit onwards, counterfactuals in event from the event onwards, and takeovers in second event from the second event onwards.

Panel a. Takeovers with less than 5 counterfactual observations with a propensity score <0.05 standard deviation are excluded. Panel b. Counterfactuals consists of pairs of firms that match the target and acquirer characteristics. Takeovers without consistent employment information in each of the three years before the merger, and takeovers with an acquirer or target for which no valid matches could be found are excluded.

[19] The number of takeovers in these estimations is somewhat smaller than for the first two estimators because we require that consistent employment information is available on both the acquirer and the target in each of the three years before the merger, and because takeovers with an acquirer or target for which no valid matches could be found drop out from the estimations.

The first column of Table 4.4 shows that takeovers have a small but significant negative impact on employment of the combined entity immediately after the merger. Employment growth in the period of the transaction is 1.86 to 2.43 percentage points lower than it would have been in the absence of a merger. The next columns reveal that the adverse effect persists for a substantial period of time. In particular, the coefficients obtained by stratified matching show that growth reductions continue to be larger than 2 percentage points in the three years after the transaction, and remain highly significant.

The coefficients of the two other estimators, which less accurately reflect how the decision to engage in a takeover depends on characteristics of both the acquirer and the target, show similar but less pronounced effects. The results based on propensity score matching, which only control for characteristics at the level of the combined entity, suggest that the negative employment effect disappears in the third year after the transaction. The coefficients based on inverse-probability-weighted regression-adjustment, which take into account that the individual characteristics of the acquirer affect the takeover decision, are more in line with the ones obtained by stratified matching. Yet, even if the stratified matching results are based on a smaller set of takeovers, the coefficient estimates are more precise than the ones based on the two other estimation approaches. This gain in precision is in line with what could be expected based on our discussion of the estimator in Section 4.3.2.

How do these workforce reductions relate to total labor force dynamics? Elsewhere we have shown that the annual gross job destruction rate in the Belgian private sector in the period of observation is 6.03 percent, compared to a job creation rate of 7.06 percent (Geurts 2016). Since takeovers represent about 6 percent of total private employment, this means that annual job loss due to takeovers in the post-merger period represents a mere 2 to 3 percent of total gross job destruction.[20] Even without taking into account the positive consequence that the workforce reductions may be partly the result of increases in labor productivity, fear of substantial job loss following takeovers seems to be largely unfounded, at least for takeover activity between domestic firms.

[20] For small growth rates, the percentage point differences in the growth rates correspond to the percentage differences between the actual and potential employment levels in the absence of a takeover. A simple back-of-the-envelope calculation shows that, as takeovers represent 6.1% of total private employment in a given year, annual work-force reductions in the post-merger periods as shown in Table 4.4 represent about 0.14% of total employment, corresponding to 2.4% of total gross job destruction in a given year.

Table 4.5 Effect of takeovers on employment growth measured over *n*-year periods
Dependent variable: employment growth rate (in percent)

	Takeover period (t-1 to t)	2-year period (t-1 to t+1)	3-year period (t-1 to t+2)	4-year period (t-1 to t+3)
Panel a.				
Propensity score matching				
	-1.86***	-3.98***	-7.26***	-6.73***
	(0.32)	(0.74)	(1.34)	(1.59)
Inverse-probability-weighted regression-adjustment				
	-2.16***	-4.83***	-7.16***	-9.11***
	(0.29)	(0.68)	(1.15)	(1.45)
No. of observations	130 274	107 789	85 642	63 767
No. of takeovers	2 249	1 929	1 517	1 063
Panel b.				
Stratified matching on acquirer and target characteristics				
	-2.43***	-4.15***	-6.22***	-6.74***
	(0.25)	(0.54)	(0.80)	(1.17)
No. of observations	5 555 003	4 439 591	3 493 527	2 577 358
No. of takeovers	1 948	1 670	1 323	932

Regression coefficients show the average treatment effects on the treated. Standard errors in parentheses

Takeovers in 2007-2012. Observations drop out from the estimations when employment is outside period of observation (after 2012). Regressions include all firms that survive or exit after year *t*. Regressions do not include counterfactuals in event from event onwards, and takeovers in second event from second event onwards.

Panel a. Takeovers with less than 5 counterfactual observations with a propensity score <0.05 standard deviation are excluded.

Panel b. Counterfactuals consists of pairs of firms that match the target and acquirer characteristics. Takeovers without consistent employment information in each of the three years before the merger, and takeovers with an acquirer or target for which no valid matches could be found are excluded

The year-by-year growth estimates presented in Table 4.4 are conditional on surviving in the preceding period. Table 4.5 shows the impact on employment over a longer period, i.e. from the pre-merger employment level in t-1 to the nth year after the transaction. The results confirm a persistent and strongly negative impact of takeover activity on firm employment growth. In the first year after the merger, employment relative to the pre-merger level is reduced by about 4 percent. Measured over the 3-year post-merger period, the adverse effect on employment growth amounts to about -7 percentage points. The results for the three estimators are broadly comparable, with the stratified matching specifications again producing more precise coefficient estimates.

Estimations of the difference in performance before and after the merger confirm the adverse effect of takeovers on employment growth. Table 4.A.3 in the Appendix presents coefficient estimates of the impact of takeovers on the difference between firm-level employment growth before and after the merger. The first column compares employment growth in the year of the takeover with that in the year before the merger. The second and third columns present results based on a comparison of two- and three-year growth rates, respectively. The negative coefficients indicate that growth of firms involved in a takeover significantly slows down after the merger and that this adverse effect persists for several periods. In the merger period, average growth rates are 0.89 percentage points lower than if the firms would not have engaged in a takeover. The negative effect amounts to -4.11 percentage points if three-year growth rates before and after the merger are compared.

We have discussed before that our estimates take into account additional changes in the firm structure that occur in the post-merger periods. Such changes have largely been neglected in previous studies. We noted that 34 percent of the merged companies are involved in another restructuring event in the 3-year period after the takeover. Misinterpreting additional acquisitions and split-offs as internal job creation and destruction, and misreading changes in the firm ID code as exits, would artificially inflate or deflate the employment growth figures. Appendix Tables 4.A.4 and 4.A.5 show coefficient estimates based on a straightforward reading of the firm employment figures in the raw data. They reveal that disregarding additional restructuring leads to a substantial underestimation of the post-merger growth performance. Year-by-year estimates in this naive approach would suggest employment growth reductions in the post-merger periods that are about twice as large as the ones presented in Table 4.4. Accumulated over the entire 3-year post-merger period, they would suggest an average employment growth reduction of 13 percentages points, instead of the -9 percentage points we find in our sample.

The estimated coefficients discussed so far represent the difference in employment growth between takeovers and their predicted outcomes in the absence of a merger. Constructing counterfactual growth rates over a relatively long period is obviously a hypothetical exercise. Moreover, the limited information in our dataset does not allow us to take into account intermediate changes in output, wages, prices, and the financial performance of the firm which affect the firm-level labor demand in a given period. The reasons behind the negative effect of takeovers on employment growth thus remain unclear. It may be the result of more efficient labor usage for the production of a constant output level, or of a reduction in output growth. Conyon et al. (2002) have shown evidence for the UK that mergers between large companies are followed by both output falls and increased labor efficiency in the merged entity. What our results do suggest, however, is that workforce reductions following takeovers are not restricted to the period of the transaction only. If the negative employment impact reflects the hypothesis that restructurings provide an opportunity for managers to shed excess labor, cut overlapping functions, and adjust the use of labor to a new optimal employment level, our results suggest that these workforce adjustments are made slowly. Similarly, if employment reductions reflect output falls, they extend over a substantial period of time.

4.5.3 Differential effect by type of takeover

To further analyze the impact of takeovers on employment, we disaggregate the sample into takeovers by small and large acquirers at the time before the transaction (*t*-1). We show results of the stratified matching estimator because of its properties of robustness and flexibility.

The estimates in Table 4.6 reveal that the negative impact of takeovers on employment growth we observed previously is for a large part attributable to workforce reductions following takeovers by small acquirers (less than 100 employees). The coefficients indicate substantial employment growth declines for this subset ranging from 2.7 to 3.6 percentage points in the year of takeover and the three post-merger periods. By contrast, takeovers by large acquirers lead to much smaller employment declines in the period of the takeover, and an unclear employment growth path in the next three periods, which suggests considerable variation in the growth performance of takeovers by large acquirers. This finding confirms Gugler at al. (2003), who showed considerable variation in post-merger performance among mergers by large companies. They found that many mergers reduce output (measured as sales) and increase efficiency, but that a large proportion of firms also increase output combined with either increases or decreases in efficiency. The size-related differences in employment growth we observe is also in line with the results for listed companies reported by Conyon et al. (2002). They found that smaller acquirers tend to exhibit greater labor demand falls and increases in labor efficiency than their larger counterparts.

Table 4.6 Effect of takeovers on employment growth by size of the acquirer in t-1
Dependent variable: employment growth rate (in percent)

	Takeover period	*n*th post-merger period		
		1st	2nd	3rd
	(t-1 to t)	(t to t+1)	(t+1 to t+2)	(t+2 to t+3)
Small acquirer (<100 employees)	-2.78***	-2.69***	-3.13***	-3.64***
	(0.30)	(0.55)	(0.65)	(0.80)
Large acquirer (≥ 100 employees)	-1.33***	-0.61	-1.89***	1.84**
	(0.35)	(0.55)	(0.64)	(0.74)

Regression coefficients show the average treatment effects on the treated using stratified matching on acquirer and target characteristics. Standard errors in parentheses.

Takeovers in 2007-2012. Observations drop out from the estimations when employment is outside period of observation (after 2012). Regressions include all firms that survive or exit in a given period. Regressions do not include firms that have exited from the year after exit onwards, counterfactuals in event from the event onwards, and takeovers in second event from the second event onwards.

Counterfactuals consists of pairs of firms that match the target and acquirer characteristics. Takeovers without consistent employment information in each of the three years before the merger, and takeovers with an acquirer or target for which no valid matches could be found are excluded.

In our previous stratified matching estimations, we controlled for the different impact of takeovers by industry by matching targets and acquirers with counterfactual firms in the same Nace 3-digit industry. To throw more light on the industry differences, Table 4.7 shows the impact of takeover activity for Manufacturing, Construction, Wholesale and retail, and Business and household services separately. The table present results of stratified matching estimations and distinguishes takeovers according to the industry of the acquiring firm. The first column of the table shows that Business and household services exhibit the strongest employment growth declines brought about by takeovers, followed by Construction and Wholesale and retail. The negative impact of takeover activity in Manufacturing is much smaller. Manufacturing, however, is dominated by larger firms, while the average firm size in other sectors, in particular in Construction and Wholesale and retail, is much smaller. Given the critical importance of the acquirer's size discussed above, we estimate the differential impact by small and large acquirers for each industry group (column 2 and 3). Two patterns emerge. On the one hand, in each industry, small acquirers lead to more substantial workforce reductions than large acquirers. This means that the firm size composition does provide some explanation for the observed industry differences. On the other hand, part of the main industry effect remains present for both small and large acquirers, with Construction and Business and household services exhibiting the strongest decreases in employment growth following a merger.

Table 4.7 Effect of takeovers on employment growth by industry and size of acquirer
Dependent variable: employment growth rate in takeover period (t-1 to t) (in percent)

	Total	Small acquirer	Large acquirer
Manufacturing	-1.45***	-1.46**	-1.38**
	(0.47)	(0.63)	(0.54)
Construction	-2.78***	-3.55***	-1.71
	(0.87)	(1.03)	(1.11)
Wholesale and retail	-1.89***	-2.35***	0.20
	(0.43)	(0.50)	(0.65)
Business and household services	-3.29***	-3.52***	-2.41***
	(0.44)	(0.51)	(0.68)

Regression coefficients show the average treatment effects on the treated using stratified matching on acquirer and target characteristics. Standard errors in parentheses.
Takeovers in 2007-2012. Counterfactuals consists of pairs of firms that match the target and acquirer characteristics.

While the size of the acquirer is found to be a major discriminating feature, the effect of takeovers on employment may also differ by other characteristics of the target and the acquirer. To further analyze these firm-level determinants, we additionally distinguish between small and large target shares, high-growth versus other targets, and related versus unrelated mergers. As explained in section 4.3.3, we use a minimum distance estimator to investigate the impact of these features.[21] Table 4.8 reports the results.

The reference class in these estimations consists of takeovers by small acquirers, where the target represents a large target share, exhibits low growth before the merger, and is in a related industry. These subsets are defined as before. The constant term in panel b. shows that the employment growth impact for this reference class is -4.1 percent in the takeover period, and amounts to -14 percent when growth is measured over the 4-year period after the transaction. The results in the first row of each panel a. and b. confirm our previous finding that the size of the acquirer is a strong determinant of post-merger employment growth.

[21] In order to have sufficient observations in each subset of takeovers, the distinction between small and large acquirers is redefined using a threshold of 50 employees.

Controlling for all other features, takeovers by large acquirers have a more positive impact on employment growth than those by small acquirers. The effect is significant up till the 3rd post-merger period. In the period of the takeover, however, the significance of the size effect disappears, and is captured by two other characteristics.

Table 4.1 showed that small acquirers tend to merge with relatively large targets, while large acquirers take over relatively smaller firms. When two firms of similar size are merged, the potential for workforce rationalizations may be greater than in the case a relatively large firm absorbs a much smaller target. The results in the second row of Table 4.8 show that the relative size of the target indeed provides some explanation for the differential effect of takeovers by small and large acquirers. As expected, the employment decline in the period of the transaction is less pronounced when relatively small targets are acquired. The effect of the target share is, however, restricted in time and disappears from the first post-merger period onwards. After that, the size of the acquirer remains the dominant explanation for differential employment growth rates following a takeover.

Takeovers targeted at incorporating new technologies of high-growth firms are expected to have a more positive impact on employment growth of the merged entity than other takeovers. Our results suggest some evidence for this. The third row of Table 4.8 shows that takeovers of targets that exhibited an average annual growth rate of at least 8 percent in the 3 years before the transaction have a more positive effect on employment growth. Panel a. indicates a significant positive effect in the period of the transaction only, but when measured over a longer period (panel b.), the accumulated effect of high growth targets remains significantly positive over a 3-year period.

Table 4.8 Effect of takeovers on employment growth by characteristics of target and acquirer

Dependent variable: employment growth rate (in percent)

Panel a.	Takeover period (t-1 to t)	*n*th post-merger period 1st (t to t+1)	2nd (t+1 to t+2)	3rd (t+2 to t+3)
Large acquirer	0.73	2.67**	2.15*	2.49**
	(0.56)	(1.08)	(1.05)	(0.89)
Small target share	1.76***	0.93	0.17	-0.86
	(0.54)	(0.89)	(0.92)	(0.79)
High growth target	1.38**	1.14	0.90	0.33
	(0.62)	(0.90)	(1.04)	(0.85)
Unrelated merger	0.13	0.63	0.27	0.22
	(0.50)	(0.86)	(0.94)	(0.75)
Constant	-4.14***	-5.21***	-4.96***	-3.44***
	(0.53)	(1.06)	(1.04)	(0.95)

Panel b.	Takeover period (t-1 to t)	2-year period (t-1 to t+1)	3-year period (t-1 to t+2)	4-year period (t-1 to t+3)
Large acquirer	0.73	3.38**	5.26**	5.90***
	(0.56)	(1.23)	(1.72)	(1.85)
Small target share	1.76***	2.82**	1.69	0.83
	(0.54)	(1.12)	(1.59)	(1.74)
High growth target	1.38**	2.90**	3.70*	2.33
	(0.62)	(1.16)	(1.80)	(1.97)
Unrelated merger	0.13	0.75	0.18	1.25
	(0.50)	(1.04)	(1.53)	(1.69)
Constant	-4.14***	-9.44***	-12.52***	-14.07***
	(0.53)	(1.23)	(1.67)	(1.86)

Regression coefficients show results of minimum distance estimations. Standard errors in parentheses.

The dummy for large acquirer' indicates takeovers by acquirers with more than 50 employees in the year before the takeover (t-1). The dummy for 'small target share' indicates takeovers where the target represents less than 25% of employment of the combined entity in t-1. The dummy for 'high growth target' indicates takeovers of targets with an average annual growth rate of more than 8% in the three years before the takeover. The dummy for 'unrelated merger' indicates mergers between firms in different industries at Nace 4-digit level (513 sectors).

The last row of Table 4.8 shows the differential effect of takeover activity by the relatedness of the acquirer and the target, as measured by their industrial classification. Conyon et al. (2002) find larger reductions in the use of labor post-merger in related mergers than in unrelated mergers and suggest that this might reflect the differing scope for rationalizations. Gugler and Yurtoglu (2004), however, find contrasting results for different countries. Our results seem to support the argument of Atalay, Hortaçsu and Syverson (2014) that vertical and horizontal expansion do not fundamentally differ. Above we showed that related and unrelated takeovers do not substantially differ with respect to the average absolute and relative sizes of the firms involved. Table 4.8 further suggests no evidence to distinguish between the employment effects of both types of takeovers. Conditional on all other characteristics, takeover activity leads to similar employment outcomes for both unrelated and related mergers.

4.6 Conclusion

This paper has provided an empirical analysis of the impact of domestic mergers and takeovers on firm employment growth across a comprehensive sample of Belgian takeovers in 2007-2012. We show how the individual characteristics of the acquiring and the acquired firm affect the selection into a takeover and subsequent employment growth. We find the size of the acquirer to be a major determinant of differential post-merger employment performance. In contrast to results that have been reported for acquisitions by listed firms in Europe, we find that takeovers by large companies (more than 100 employees) do not lead to systematic decreases in employment growth following in the merger. Instead, the merged entities exhibit substantial variation in post-merger employment growth, which is present across various industries and subsets of mergers. This suggests that workforce rationalizations are not the dominant motivation for takeover activity by larger firms. A wide range of other motivations, as suggested in the merger literature, may drive the decision to takeover, and lead to either increases or decreases in employment. Takeovers targeted at acquiring new technologies are one example of proposed synergy gains that may drive takeover activity. Our results suggest some evidence for this as takeovers of high-growth targets have a more positive impact on employment growth than other takeovers.

By contrast, our results suggest strong evidence for significant employment decreases following takeovers by small acquirers (less than 100 employees). In the period of the transaction, employment growth is about 3 percentage points lower than it would have been in the absence of a merger, and this adverse effect persists for three post-merger years. The negative employment impact is significant after controlling for various characteristics of the target firm, and is observed in different industries. If the employment declines reflect important workforce rationalizations, our results suggest that these adjustments are made slowly and extend over a substantial period of time.

In the present version of the paper, we have provided a general picture of how different types of acquirers and targets lead to differential employment outcomes. Our setting, which relies on large subsets of counterfactual observations, however, enables us to further investigate detailed subsets of the sample which can be more explicitly linked to different motivations for takeover activity. Mergers in industries with declining demand, for example, are likely to be motivated by rationalizations of capacity; acquisitions in industries that exhibited an important wave of consolidation can be linked to oligopolistic behavior; mergers in industries with high levels of unionization are likely to aim at productivity gains by shedding excess labor; takeovers targeted at young high-growth companies in innovative industries can be linked to technology acquisitions; or takeovers by large companies targeted at suppliers in related industries can be linked to vertical integration along the value chain. In this sense, the research presented in this doctoral thesis is not finished. In fact, it has only begun.

Appendix

Table 4.A. 1 Description of the variables

Variables		Description
E_{it}	Employment of firm *i* at time *t*	The number of employees firm *i* employs at June 30 of year *t*.
g_{it}	Employment growth rate of firm *i* in period *t*-1 to *t*	The growth rate equals $g_{it} = (E_{it} - E_{it-1})/\bar{E}_{it}$, with $\bar{E}_{it} = (E_{it} + E_{it-1})/2$.
M_{it}	Dummy variable for takeover in period *t*-1 to *t*	Takes a value of one if firm *i* is involved in a takeover in period *t*-1 to *t.*
C_t	Industry concentration ratio at time *t*	Employment share of the four largest firms at the Nace 3-digit industry level (166 industries) at time *t*
hi_t	Industry concentration index at time *t*	Herfindahl index at the Nace 3-digit industry level measured as the sum of the squares of the employment shares of the firms within the industry at time *t*
PA_i	Parent firm dummy	Takes a value of one if firm *i* owns at least 50 percent of the shares of another Belgian firm in the sample period (2004-2012)
G_i	Enterprise group dummy	Takes a value of one if firm *i* either controls or is controlled by another Belgian firm in the sample period (2004-2012)
F_i	FDI dummy	Takes a value of one if firm *i* is either receiver of sender of FDI in the sample period (2004-2012)
Subsets		
Small versus large acquirer		A small acquirer has less than 100 employees at the time before the takeover (*t*-1). Otherwise the acquirer is denoted as large. Table 4.8: a small acquirer has less than 50 employees at the time before the takeover (*t*-1).
Small versus large target share		A small target share denotes takeovers where the employment of the acquired firm represents less than 25 percent of the employment of the combined entity at the time before the takeover (*t*-1). Otherwise the target share is denoted as large.
High growth target		A high-growth target is an acquired firms with an average annual growth rate of at least 8 percent in the three periods before the merger (*t*-3, *t*-2 and *t*-1).
Related versus unrelated takeovers		Takeovers are classified into related and unrelated depending on whether the target(s) and the acquirer belonged to the same Nace 3-digit industry (166 industries).

Table 4.A. 2 Share of firms involved in a restructuring event

	Obs.	Share in event		
	(*t*)	(*t* to *t*+1)	(*t* to *t*+2)	(*t* to *t*+3)
Takeovers	2 259	0.14	0.25	0.34
Firms in control sample	128 025	0.02	0.03	0.05

Table 4.A. 3 Effect of takeovers on growth acceleration

Dependent variable: difference in employment growth rate pre- and post-merger (in ppt)

	1-period growth difference	2-period growth difference	3-period growth difference
	(*t*-1 to *t*) - (*t*-2 to *t*-1)	(*t*-1 to *t*+1) - (*t*-3 to *t*-1)	(*t*-1 to *t*+2) - (*t*-4 to *t*-1)
	-0.89**	-1.82***	-4.11***
	(0.37)	(0.68)	(0.98)
No. of observations	5 555 003	4 439 591	3 464 952
No. of takeovers	1 948	1 670	1 323

Regression coefficients show the average treatment effects on the treated using stratified matching on acquirer and target characteristics. Standard errors in parentheses.

Takeovers in 2007-2012. Regressions include all firms that survive or exit in a given period. Regressions do not include counterfactuals in event from the event onwards, and takeovers in second event from the second event onwards.

Counterfactuals consists of pairs of firms that match the target and acquirer characteristics. Takeovers without consistent employment information in each of the three years before the merger, and takeovers with an acquirer or target for which no valid matches could be found are excluded.

Table 4.A. 4 Effect of takeovers on year-by-year employment growth when additional restructuring events post-merger are neglected

Dependent variable: employment growth rate (in percent)

	Takeover period	*n*th post-merger period		
		1st	2nd	3rd
	(*t*-1 to *t*)	(*t* to *t*+1)	(*t*+1 to *t*+2)	(*t*+2 to *t*+3)
Propensity score matching				
	-2.99***	-3.45***	-3.60***	-4.78***
	(0.33)	(0.93)	(1.00)	(1.18)
Inverse-probability-weighted regression-adjustment				
	-2.79***	-3.52***	-4.04***	-5.24***
	(0.30)	(0.88)	(0.94)	(1.11)
No. of observations	134 755	134 755	132 100	108 315
No. of takeovers	2 151	2 148	2 086	1 642

Regression coefficients show the average treatment effects on the treated. Standard errors in parentheses. Takeovers with less than 5 counterfactual observations with a propensity score <0.05 standard deviation are excluded. Regressions include all firms that survive or exit in a given period. Regressions do not include firms that have exited from the year after exit onwards.

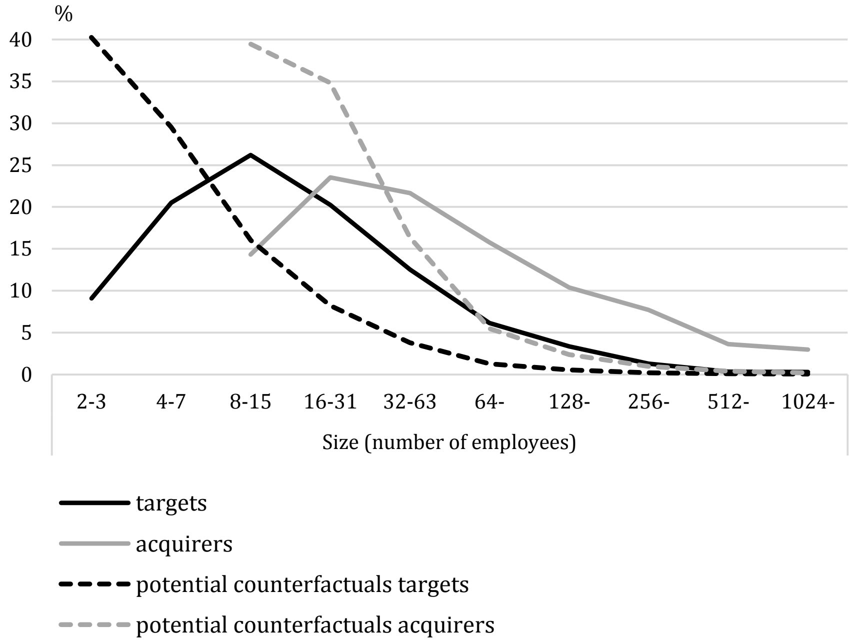
Table 4.A. 5 Effect of takeovers on *n*-year period employment growth when additional restructuring events post-merger are neglected

Dependent variable: employment growth rate (in percent)

	Takeover period	2-year period	3-year period	4-year period
	(*t*-1 to *t*)	(*t*-1 to *t*+1)	(*t*-1 to *t*+2)	(*t*-1 to *t*+3)
Propensity score matching				
	-2.99***	-6.15***	-9.55***	-12.42***
	(0.33)	(0.97)	(1.34)	(1.80)
Inverse-probability-weighted regression-adjustment				
	-2.79***	-6.21***	-9.80***	-13.45***
	(0.30)	(0.93)	(1.26)	(1.68)
No. of observations	134 755	134 755	134 755	113 091
No. of takeovers	2 151	2 151	2 151	1 750

Regression coefficients show the average treatment effects on the treated. Standard errors in parentheses. Takeovers with less than 5 counterfactual observations with a propensity score <0.05 standard deviation are excluded. Regressions include firms that exit after year *t*. Regressions do not include observations outside period 2003-2012.

Figure 4.A. 1 Size distribution of targets and acquirers and their potential counterfactuals at the time before the takeover (t-1)



References

Abbring, J.H. and J.R. Campbell (2005). "A firm's first year." Tinbergen Institute Discussion Paper No. 2005-046/3.

Abowd, J.M., and F. Kramarz (1999). "The Analysis of Labor Markets Using Matched Employer-Employee Data." In O.C. Ashenfelter and D. Card (eds.), Handbook of Labor Economics Volume 3, Part B (pp. 2629–2710). Elsevier.

Abowd, J.M., P. Corbel and F. Kramarz (1999). The entry and exit of workers and the growth of employment: an analysis of French establishments. *The Review of Economics and Statistics*, 81:2, 170-187.

Acs, Z. J. and C. Armington (1998). "Longitudinal Establishment and Enterprise Microdata (LEEM) Documentation." CES 98–9, Center for Economic Studies, US Bureau of the Census.

Acs, Z.J., T. Astebro, D. Audretsch and D.T. Robinson (2016). Public policy to promote entrepreneurship: a call to arms. *Small Business Economics*, 47:1, 35-51.

Acs, Z. J. and D.B. Audretsch (1989). Births and Firm Size. *Southern Economic Journal*, 56:2, 467-475.

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

Almus, M. and E.A. Nerlinger (2000). Testing "Gibrat's Law" for young firms – Empirical results for West Germany. *Small Business Economics*, 15, 1-12.

Anderson, D.W., L. Kish and R.G. Cornell (1980). On Stratification, Grouping and Matching. *Scandinavian Journal of Statistics*, 7:2, 61-66.

Angrist, J.D. and J.S. Pischke (2009). "Mostly Harmless Econometrics. An Empiricist's Companion." Princeton University Press.

Asker, J., A. Collard-Wexler and J. De Loecker (2014). Dynamic inputs and resource (mis)allocation. *Journal of Political Economy*, 122:5, 1013-1063.

Atalay, E., A. Hortaçsu and C. Syverson (2014). Vertical integration and input flows. *American Economic Review*, 104:4, 1120-1148.

Audretsch, D.B., E. Santarelli and M. Vivarelli (1999). Start-up size and industrial dynamics: Some evidence from Italian manufacturing. *International Journal of Industrial Organization*, 17, 965-983.

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. and G. Picot (1995). Employment generation by small producers in the Canadian manufacturing sector. *Small Business Economics*, 7:4, 317-331.

Baldwin, J.R. and P.K. Gorecki (1987). Plant Creation versus Plant Acquisition. The Entry Process in Canadian Manufacturing. *International Journal of Industrial Organization*, 5:1, 27-41.

Bartelsman, E., S. Scarpetta and F. Schivardi (2005). Comparative analysis of firm demographics and survival: Evidence from micro-level sources in OECD countries. *Industrial and Corporate Change*, 14:3, 365-391.

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

Beck, T., A. Demirgüç-Kunt, L. Laeven and V. Maksimovic (2006). The determinants of financing obstacles. *Journal of International Money and Finance*, 25, 932-952.

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.

Bhagat, S., A. Shleifer, R.W. Vishny, G. Jarrel and L. Summers (1990). "Hostile takeovers in the 1980s: the return to corporate specialization." Brookings Papers on Economic Activity: Microeconomics, Vol. 1990, 1–84.

Bilsen, V. and J. Konings (1998). Job creation, job destruction, and growth of newly established, privatized, and state-owned enterprises in transition economies: survey evidence from Bulgaria, Hungary, and Romania. *Journal of Comparative Economics*, 26:3, 429–445.

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

Brockhaus, R.H. (1980). Risk taking propensity of entrepreneurs. *Academy of Management Journal*, 23:3, 509-520.

Brown, C. and J.L. Medoff (1988). "The impact of firm acquisitions on labor." In A.J. Auerbach (ed.), Corporate takeovers: Causes and consequences. University of Chicago Press.

Brown, J.D., J.S. Earle and Y. Morgulis (2015). "Job creation, small vs. large vs. young, and the SBA." NBER Working Paper No. 21733.

Butani, S.J., R.L. Clayton, V. Kapani, J. Spletzer, D.M. Talan and G.S. Werking (2006). Business employment dynamics: Tabulations by employer size. *Monthly Labor Review*, 2:1, 3-22.

Bycroft, C (2003). "Record linkage in LEED: An Overview." Statistics New Zealand.

Cabral, L. (1995). Sunk costs, firm size and firm growth. *Journal of Industrial Economics*, 43:2, 161-172.

Cabral, L. 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*, 1:2, 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.

Caves, R.E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, 36:4, 1947-1982.

Caves, R.E. and M.E. Porter (1977). From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition. *Quarterly Journal of Economics*, 91:2, 241-262.

Clayton, R.L. and J.R. Spletzer (2009). "Business Employment Dynamics." In T. Dunne, J.B. Jensen and M.J. Roberts (eds.), Producer Dynamics: New Evidence from Micro Data (pp. 125-147). University of Chicago Press.

Coase, R.H. (1937). The nature of the firm. *Economica*, 4:16, 386–405.

Contini, B., R. Leombruni, L. Pacelli and C. Villosio (2007). "Wage Mobility and Dynamics in Italy in the 90's." NBER Working Paper No. w13029.

Conyon, M.J., S. Girma, S. Thompson and P.W. Wright (2001). Do hostile mergers destroy jobs? *Journal of Economic Behavior & Organization*, 45, 427-440.

Conyon, M.J., S. Girma, S. Thompson and P.W. Wright (2002). The impact of mergers and acquisitions on company employment in the United Kingdom. *European Economic Review*, 46:1, 31–49.

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.C. Haltiwanger and S. Schuh (1996a). "Job creation and destruction." Cambridge: MIT Press.

Davis, S., 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.

de Wit, G. and J. de Kok (2014). Do small businesses create more jobs? New evidence for Europe. *Small Business Economics*, 42:2, 283-295.

Decker, R., J.C. Haltiwanger, R. Jarmin and J. Miranda (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, 28:3, 3-24.

Dixon, J. and A.-M. Rollin (2012). "Firm Dynamics: Employment Growth Rates of Small Versus Large Firms in Canada." Ottawa: Statistics Canada.

Dunne, T. and A. Hughes (1994). Age, size, growth and survival: UK companies in the 1980s. *Journal of Industrial Economics*, 42:2, 115-140.

Dunne, T., M.J. Roberts and L. Samuelson (1988). Patterns of firm entry and exit in U.S. manufacturing industries. *RAND Journal of Economics*, 19:4, 495-515.

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.

Dutz, M.A. (1989). Horizontal mergers in declining industries: Theory and evidence. *International Journal of Industrial Organization*, 7, 11-33.

Ericson, R. and A. Pakes (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62:1, 53-82.

Eriksson, T. and J.M. Kuhn (2006). Firm spin-offs in Denmark 1981–2000. Patterns of entry and exit. *International Journal of Industrial Organization*, 24:5, 1021–1040.

European Commission (2013). "The Entrepreneurship 2020 Action Plan. Reigniting the entrepreneurial spirit in Europe."

Eurostat-OECD (2007). "Eurostat-OECD Manual on Business Demography Statistics." Luxembourg: European Communities and OECD.

Evans, D.S. and B. Jovanovic (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy*, 97:4, 808-827.

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.

Evans, D.S. and L.S. Leighton (1989). Some empirical aspects of entrepreneurship. *American Economic Review*, 79:3, 519-535.

Fazzari, S, R.G. Hubbard and B. Petersen (1988). "Financing constraints and corporate investment." Brookings Papers on Economic Activity, 1, 141-206.

Foster L., J.C. Haltiwanger and C.J. Krizan (2001). "Aggregate Productivity Growth. Lessons from Microeconomic Evidence." In C.R. Hulten, E.R. Dean and M.J. Harper (eds.), New Developments in Productivity Analysis (pp. 303-372). University of Chicago Press.

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

Fritsch M. and P. Mueller (2004). Effects of New Business Formation on Regional Development over Time. *Regional Studies*, 38:8, 961–975.

Geroski, P.A. (1995). What do we know about entry? *International Journal of Industrial Organization*, 13:4, 421–440.

Geurts, K. (2016). Longitudinal firm-level data: problems and solutions, *Small Business Economics*. 46:3, 425-445.

Geurts, K. and J. Van Biesebroeck (2014). "Job creation, firm creation, and *de novo* entry." CEPR Discussion Paper No. 10118.

Geurts, K., M. Ramioul and P. Vets (2009). "Employee flows to improve measures of job creation and destruction and of firm dynamics. The case of Belgium." HIVA Working Paper, KU Leuven.

Girma, S. and H. Görg (2003). "Blessing or Curse? Domestic Plants' Survival and Employment Prospects after Foreign Acquisitions." IZA Discussion Paper No. 706.

Grossman, S. and O.D. Hart (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94:4, 691–719.

Gugler, K. and B. Yurtoglu (2004). The effects of mergers on company employment in the USA and Europe. *International Journal of Industrial Organization*, 22:4, 481–502.

Gugler, K., D.C. Mueller, B.B. Yurtoglu and C. Zulehner (2003). The effects of mergers: An international comparison. *International Journal of Industrial Organization*, 21:5, 625–653.

Guner, N., G. Ventura and Y. Xi (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11, 721-744.

Hall, B.H. (1987). The relationship between firm size and firm growth in the U.S. manufacturing sector. *Journal of Industrial Economics*, 35:4, 583-606.

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

Helfat, C.E. and M.B. Lieberman (2002). The birth of capabilities: market entry and the importance of pre-history. *Industrial and Corporate Change*, 11:4, 725-760.

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.

Holmes, T.J. and J.A. Schmitz (2010). "Competition and productivity: A review of evidence." Federal Reserve Bank of Minneapolis: Research Department Staff Report 439.

Holmström, B. and J. Roberts (1998). The Boundaries of the Firm Revisited. *The Journal of Economic Perspectives*, 12:4, 73-94.

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

Hsieh, C.-T. and P.J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124:4, 1403-1448.

Imbens, G.W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86:1, 4–29.

Imbens, G.W. and J.M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47:1, 5-86.

Jarmin, R.S. and J. Miranda (2002). "The Longitudinal Business Database." Center for Economic Studies Discussion Paper No. CES-WP-02-17.

Jarrell, G.A., J.A. Brickley and J.M. Netter (1988). The market for corporate control: The empirical evidence since 1980. *Journal of Economic Perspectives*, 2:1, 49-68.

Jensen, M.C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review*, 76:2, 323-329.

Jensen, M.C. (1988). Takeovers: Their causes and consequences. *Journal of Economic Perspectives*, 2:1, 21-48.

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

Kirchhoff, B.A. and P. Greene (1998). Understanding the theoretical and empirical content of critiques of U.S. job creation research. *Small Business Economics*, 10, 153–169.

Konings, J. (1995). Gross job flows and the evolution of size in UK establishments. *Small Business Economics*, 7:3, 213-220.

Konings, J., H. Lehmann and M. Schaffer (1996). Job creation and job destruction in a transition economy: Ownership, firm size, and gross job flows in Polish manufacturing 1988-91. *Labour Economics*, 3:3, 299-317.

Korkeamäki, O. and T. Kyyrä (2000). "Integrated panel of Finnish companies and workers." VATT Discussion Paper 226. Helsinki: Government Institute for Economic Research.

Lafontaine, F. and M. Slade (2007). Vertical Integration and Firm Boundaries: The Evidence. *Journal of Economic Literature*, 45:3, 629-685.

Lehto, E. and P. Böckerman (2008). Analysing the employment effects of mergers and acquisitions. *Journal of Economic Behavior & Organization*, 68:1, 112–124.

Lichtenberg, F.R. and D. Siegel (1989). „The effect of takeovers on the employment and wages of central-office and other personnel." NBER Working Paper No. 2895.

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.

Lucas, R.E. (1978). On the size distribution of business firms. *Bell Journal of Economics*, 9:2, 508-523.

Mansfield, E. (1962). Entry, Gibrat's Law, innovation, and the growth of firms. *American Economic Review*, 52:5, 1023-1051.

Margolis, D.N. (2006). Should employment authorities worry about mergers and acquisitions? *Portuguese Economic Journal*, 5, 167–194.

Mata, J. (1993). Entry and type of entrant. Evidence from Portugal. *International Journal of Industrial Organization*, 11:1, 101-122.

Mata, J. (1994). Firm 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 Organization*, 13:4, 459-481.

McGuckin, R.H. and S.V. Nguyen (2001). The impact of ownership changes: a view from labor markets. *International Journal of Industrial Organization*, 19, 739-762.

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

Mitchell, M.L. and J.H. Mulherin (1996). The impact of industry shocks on takeover and restructuring activity. *Journal of Financial Economics*, 41, 193-229.

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 (2015). “Entrepreneurship at a Glance 2015.” Paris: OECD Publishing.

Palepu, K.G. (1986). Predicting takeover targets: A methodological and empirical analysis. *Journal of Accounting and Economics*, 8, 3-35.

Persson, H. (1998). “Job flows and worker flows in Sweden 1986-1995.” In H. Persson, Essays on Labour Demand and Career Mobility (pp. 1-84). Stockholm: Swedish Institute for Social Research.

Persson, H. (2004). The Survival and Growth of New Establishments in Sweden, 1987-1995. *Small Business Economics*, 23:5, 423-440.

Pesola, H. (2009). “Labour Market Transitions Following Foreign Acquisitions.” Helsinki: HECER Discussion Paper No. 251.

Pinkston, J.C. and J.R. Spletzer (2002). “Annual measures of job creation and job destruction created from quarterly microdata.” Washington: BLS Statistical Survey Papers.

Prais, S.J. (1958). The statistical conditions for a change in business concentration. *Review of Economics and Statistics*, 40:3, 268-272.

Reynolds, P., D.J. Storey and P. Westhead (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies*, 28:4, 443-456.

Robertson, K., L. Huff, G. Mikkelson, T. Pivetz and A. Winkler (1999). “Improvements in record linkage processes for the Bureau of Labor Statistics’ Business Establishment list.” In Record Linkage Techniques—1997: Proceedings of an International Workshop and Exposition, 212–221. Washington, D.C.: National Academy Press.

Roll, R. (1986). The Hubris Hypothesis of Corporate Takeovers. *Journal of Business*, 59:2, 197-216.

Schumpeter, J.A. (1942). “Capitalism, Socialism and Democracy.” New York: Harper.

Shapiro, C. (1989). “Theories of oligopoly behavior.” In R. Schmalensee and R. Willig (eds.), Handbook of Industrial Organization. Vol. 1 (pp. 329–414). Amsterdam.

Shleifer, A. and L.H. Summers (1988). “Breach of trust in hostile takeovers.” In A.J. Auerbach (ed.), Corporate takeovers: Causes and consequences (pp. 33–68). University of Chicago Press.

Siegfried, J.J. and L.B. Evans (1994). Empirical studies of entry and exit: A survey of the evidence. *Review of Industrial Organization*, 9:2, 121-155.

Storey, D.J. (1991). The birth of new firms. Does unemployment matter? A review of the evidence. *Small Business Economics*, 3:3, 167-178.

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

Vilhuber, L. (2008). “Adjusting Imperfect Data: Overview and Case Studies.” In E.P. Lazear and K.L. Shaw (eds.), The Structure of Wages: An International Comparison (pp. 59-80). University of Chicago Press.

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

Wagner, J. (2007). Exports and Productivity: A Survey of the Evidence from Firm-level Data. *The World Economy*, 30:1, 60-82.

Williamson, O.E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics*, 22:2, 233-261.

Wooldridge, J.M. (2010). “Econometric Analysis of Cross Section and Panel Data.” Second Edition. Cambridge: MIT Press.