Toward an integrated approach to industry dynamics and labor mobility

Ricardo Mamede

Industrial dynamics and labor mobility are two fields of economic research that have developed fast in recent years, but along what are essentially separate lines. This article shows that the processes those two fields deal with can be highly interdependent, and discusses the usefulness of, and the opportunity for, an integrated approach to the dynamics of industries and labor mobility. It concludes with a list of questions that inform a research agenda dedicated to such approach.

1. Introduction

Industrial dynamics and studies of labor mobility are two fields of economic research that have developed fast in the past two decades. In both cases, such development was very much related with the increased availability of micro data, of computational resources, and of statistic and econometric tools suitable to their treatment. These, in turn, have favored the identification of a number of empirical regularities (which are often taken as “stylized facts” in both domains). On the basis of such evidence, existing theoretical models were tested and new models were developed aiming at a better explanation of the regularities found in the data.

A further common feature of those two fields of research is the fact that both deal with what can be seen as epiphenomena of the dynamic nature of the contemporary capitalist societies. We now know that the turbulence in industry structures—as a result of entry and exit of firms, changes in market shares, changes in property control, etc.—is striking. For example, using a harmonized firm level dataset of 24 industrial and developing countries, Bartelsman et al. (2004) found that, even when micro firms (i.e., firms with less then 20 people) are excluded, the annual sum of entries and exits is between 3% and 8% of the total number of firms in most industrial countries; with micro firms included, the figure increases to 20–25%. When we look at the figures on labor market dynamics, the picture is no less impressive: according to the OECD (1999), the annual turnover of the workforce in industrial economies (understood as the sum of hires and separations) varies between 10% and 15% (whilst total employment typically does not change >1–2%).

© The Author 2008. Published by Oxford University Press on behalf of Associazione ICC. All rights reserved.
There are plenty of reasons to believe that changes in industry structure and worker mobility are not entirely independent phenomena. At the most obvious level, the growth of existing firms and the creation of new ones is necessarily related to an inflow of workers to those firms, just as the contraction and closure of firms have the opposite effects on the supply-side of the labor markets (Davis et al., 1996). Moreover, industry turbulence affects the labor markets not only in such direct way, but also indirectly through the vacancy chains that are opened and closed by firms’ growth/founding and contraction/failure (as pointed out e.g., by Haveman, 1995). Reverting the direction of causality, it has been noted for a long time (e.g., Staw, 1980) that worker turnover has both positive and negative consequences for organizations, and in this sense they may constitute an important determinant of industry dynamics.

In this article, I will argue that notwithstanding all the possible interdependencies between industry evolution and labor market dynamics, there remains a lack of systematic discussion about the details of such coupled dynamics and its implications. In fact, most theoretical models of industrial dynamics (for surveys see e.g., Dosi et al., 1997; Sutton, 1997; Caves, 1998) tend to focus on the technological or financial determinants of changes in the structure of industries, abstracting from the influence of labor market determinants. In the same vein, the reference models of worker mobility (for a survey see e.g., Farber, 1999) typically underestimate the mutual influence between industry dynamics and labor market forces. With a few notable exceptions, most of the empirical work that has been done in both fields of research has followed along the same lines.

In many contexts, ignoring the mutual influence between the evolution of industry structures and the patterns of worker mobility does not do much harm to the advance of knowledge. While it is difficult to imagine situations in which the two dynamic processes are entirely independent, it is clear that the movement of workers between firms tends to be a minor issue in the evolution of several industries (specially those that rely on low-skilled, homogeneous labor, and/or in which firms operate as monopsonists, or quasi-monopsonists, within the relevant labor markets); similarly, the movement of workers in the labor market is only partly determined by the evolution of the firms that employ them—cultural, institutional, and/or idiosyncratic factors usually exert their influence and may often be more relevant than industry turbulence in determining the patterns of worker mobility. In such contexts, abstracting from the influence of worker turnover on industry evolution, or vice-versa, simply reflects the need to concentrate on the essentials and leave aside the details, which is common to any scientific endeavor.

However, we also know that such mutual influence can be crucial in many other contexts. In fact, historical accounts of industries which are highly dependent on a specialized labor force often show that the patterns of firms’ evolution and of labor force mobility are intrinsically related. For example, in relation to both hi-tech (Baron, 2004) and professional services industries (Gallouj and Gallouj, 1996;
Mamede, 2002) it has been emphasize that the performance of firms is very much affected by their capacity to recruit skilled workers and to avoid poaching by competitors. Such “recruitment-based competition” (to use the expression suggested by Sørensen, 2004), together with the highly turbulent character of some of those industries (especially those in the early phases of their life-cycle), also imply that the movement of workers will be strongly influenced by the dynamics of the relevant population of employing organizations. When this is the case, theoretical and empirical inquiries of industry evolution, which abstract from the role of labor market dynamics—or vice-versa—risk missing the main elements of the dynamic picture they propose to explain.

It is thus worthwhile to look at where we stand in our knowledge of the interdependencies between the evolution of industry structures and the patterns of worker mobility between firms, to signal the gaps in the relevant literature, and to point toward possible developments that may help us elucidate the dynamic processes involved. These constitute the central aims of this article.

The remaining sections are organized as follows. Sections 2 and 3 are dedicated to the separate analysis of two opposite influences: first, I discuss the extent to which the empirical literature has considered the role of firms and industries in explaining the turnover of workers, and then I turn to the analysis of the research efforts related to the impact of worker turnover on the evolution of industries. Section 4 discusses how the theoretical literature has dealt with the coupled dynamics at hand, and suggests different possible strategies to put industry and labor market dynamics together in an integrated theoretical framework. Section 5 summarizes the main contributions of the article and concludes with a list of questions that may inform a research agenda in this domain.

2. Evidence on the role of firms and industries in explaining the inter-firm mobility of workers

2.1 Industry dynamics and the gross creation and destruction of jobs

The literature on job creation and destruction provides the first instance of a direct link between research on industrial dynamics and work on labor market flows. In fact, this literature—which focuses on traditional topics in labor economics—as both benefited from and contributed to the theory and evidence produced in the realm of industrial dynamics.

During the 1980s, evidence on the pervasiveness of entry and exit of firms accumulated continuously. Dunne et al. (1988), for example, have used data from the US Census Bureau, which included information collected by five “Census of Manufacturing” from 1963 to 1982, in order to study the patterns of entry and exit in US manufacturing industries. They have shown that, even disregarding the smallest
firms, 38.6% of the firms included in each census were not included in the previous one (which typically took place 5 years before). The authors have also shown that, although numerous, entrants tend to be much smaller than incumbent firms, being responsible for only 15.8% of the industry output. Similar results were obtained with respect to firm exits (with the market share of the exiting firms being slightly higher). These results corroborated the evidence already produced by the empirical literature on the so-called “Gibrat’s Law” (e.g., Evans, 1987; Hall, 1987),¹ which has shown moreover that firm growth is negatively related with not only firm size and age (with younger firms facing a higher probability of failure) but also with better growth perspectives for those that survive.

The literature on job creation and job destruction has established a link between such patterns of turbulence in industry structures and the gross flows of jobs in the labor markets. The following definitions (or minor variations of it) are central to the understanding of such link in the context of this literature (Davis and Haltiwanger, 1992): gross job creation at time $t$ corresponds to the employment gains summed over all business units that expand or start up between $t−1$ and $t$; gross job destruction at time $t$ corresponds to the employment losses summed over all business units that contract or shut down between $t−1$ and $t$; gross job reallocation at time $t$ is the sum of all business unit’s employment gains and losses that occurred between $t−1$ and $t$ (it equals the sum of gross job creation and job destruction). The corresponding rates are obtained by dividing those variables by the total employment at $t$ (or, as is often the case, by the arithmetic mean of total employment in periods $t$ and $t−1$).

This stream of literature has produced an immense amount of evidence on some crucial aspects of the labor market dynamics. Davis and Haltiwanger (1999) review the main results that were obtained by studies conducted in several different countries during the 1990s; on the basis of such studies they show that: around 10% of jobs are created and other 10% are destroyed every year; in every country, the rate of job reallocation is higher than 10% for most of the sectors at a two-digit desegregation level (using the international system of industrial classification); most of the job creation (destruction) is due to the expansion (contraction) of existing firms, rather than to firm entries (exits).

For example, using data from the US Annual Survey of Manufacture between 1972 and 1986, Davis and Haltiwanger (1992) have found that: the annual rates of job creation and job destruction at the plant level were 9.2% and 11.3%, respectively (Dunne et al., 1989, have reached similar results, using different data); entries were responsible for 20% of job creation and exits by 25% of job destruction; not only is job destruction mostly driven by the contraction of existing firms, but also about three-fourth of job destruction takes place in plants that lose more than 20% of employment in 1 year.

¹For a review of the debate surrounding the “Gibrat’s Law” of proportionate effects see Sutton (1997).
More importantly in the present context, Davis and Haltiwanger (1992) show that the inter-sectoral reallocation of jobs plays a minor role in explaining total job reallocation (this is true even if sectors are defined not only in terms of industrial classification, but also according to plants’ age, size, ownership type, or region). That is, they show that explaining total employment reallocation implies taking into account the causes underlying the simultaneous occurrence of job creation and job destruction within each sector.2

As one could expect, different industries do not show the same intensity of job reallocation. Some of the studies that have been done within this line of inquiry since the early 1990s have thus tried to identify the determinants of inter-industrial differences in the patterns of job creation and destruction. For example, using data on the Swedish economy between 1986 and 1997, Antelius and Lundberg (2003) have found that job reallocation is: lower in manufacturing industries than in services, higher in more innovative and fast growing industries, higher in industries with smaller firms and lower returns, lower in industries with more stable market shares, lower in the more export-oriented industries in which the presence of foreign capital is more pronounced. These results clearly suggest that the same kind of features that characterize the more turbulent industries—innovativeness, fast growth, competitive pressure, etc.—are also present in the explanation of higher degrees of job reallocation.

Such coincidence, however, is not at all surprising. In fact, the definitions of job creation, job destruction, and job reallocation which were presented above are indeed not indicators of worker flows, but rather indicators of net employment changes summed over all business units belonging to some category. And it is only natural that higher levels of net employment changes (in absolute terms) occur in more turbulent industries.

The use of such definitions constitutes both the crucial strength and the crucial weakness of the “job creation and job destruction” type of approach. On one hand, it has allowed the production of a considerable amount of new evidence on the heterogeneity of business units in terms of job dynamics (a line of research which had a significant impact in such diverse fields as labor economics, industrial organization, and macroeconomics), drawing on data sources that were readily available. While most evidence on labor market mobility had been previously produced using information on individuals, Davis, Haltiwanger, and others took advantage of databases dedicated to the demand-side of the labor market (actually using the same type of information used by researchers of industry dynamics) to explore this field from a different perspective.

2The authors show that, although all two-digit sectors have experienced a contraction in total employment during the period under analysis (1972–1986), in every sector there were plants in which job creation took place (the same applies up to the four-digit desegregation).
On the other hand, by looking only at the net employment changes in business units, research on job creation and destruction is unable to capture a significant part of the labor market flows—namely, it ignores all the changes in the composition of the workforce that do not lead to variations in total employment within firms. This represents an obstacle to the understanding of the mutual influence between industry dynamics and worker mobility, since it is not indifferent to a firm whether the amount of changes in the composition of its workforce has exceeded or not the number of hires/separations needed to accommodate the expansion/contraction of the firm. For example, even if the total number of employees has remained unchanged, if half of the workers left since last period and half of the personnel is new to the firm in the current period this may have a significant impact in the performance of the firm.

The type of work to be discussed in the following section has tried to overcome this limitation by looking simultaneously at both sides of the labor market in the analysis of worker mobility.

2.2 Worker turnover in excess of job reallocation and its determinants

To a large extent, the shortcomings of the “job creation and job destruction” approach discussed before reflect the absence of adequate data to carry out an integrated analysis of labor market dynamics: while the availability of longitudinal data from different countries concerning either individual workers or business units has increased sharply since the early 1980s, databases matching the trajectory of both workers and firms in different time periods—the kind of data that allow to analyze not only the net employment changes, but also the total turnover of workers at the firm level—are still scarce.

The use of matched employer–employee longitudinal databases\(^3\) provides the basis for more precise estimations concerning the relative importance of demand- and supply-side determinants in explaining the mobility of workers between productive units. In one of the first studies providing direct evidence on this issue, drawing on administrative data for eight States in the USA for the period 1978–1984, Anderson and Meyer (1994) estimate that only 31% of the quarterly total worker turnover (i.e., the sum of all hires and separations) was explained by the creation and destruction of jobs (i.e., by the demand-side of the labor market). Similar results were achieved by Hamermesh et al. (1996) in their study of the Dutch economy in the period 1988–1990 (job turnover rate was found to be 6.2%, about one-third of the figure estimated for worker turnover, 22%). Albaek and Sorensen (1998), using data from Denmark for the period 1980–1991, found that, on average, job creation

\(^3\)For an overview of the different studies that have used matched employer–employee data see Abowd and Kramarz (1999). Hamermesh (1999) discusses some research avenues (which overlap only partly with what is discussed in the present article) that are opened with the increased availability of this type of data.
constituted 42% of hiring, and job destruction represented 41% of separations. Abowd et al. (1999) used a representative sample of French establishments from 1987 to 1990 to show that annual job creation is characterized as hiring three persons and separating two for each job created in a given year, and that annual job destruction is characterized as hiring one person and separating two for each job destroyed in a given year. Tattara and Valentini (2005), using data for the Veneto region (Italy) for the period 1982–1996, estimate that, on average, job creation constitutes 36% of total accessions and job destruction makes up 34% of separations. Finally, using quarterly data for the State of Maryland in the USA during the period 1985–1994, Burgess et al. (2000a) found that job flows account for nearly 30% of the worker flows in nonmanufacturing industries, and about 37.6% in manufactures.4

All these studies demonstrate that underneath the net changes in total employment at the firm level, there is a considerable amount of simultaneous hiring and separations going on. The fact that the turnover of workers is only partly determined by the expansion and contraction of business units should not, however, be taken as an indication of little mutual influence between labor market flows and industry dynamics. While there are a number of different factors which may explain worker flows in excess of job reallocation—or “churning,”5 as Lane, Stevens, and colleagues (see references below) call it—it is plausible that at least part of those factors are closely related with the dynamics of firms and industries.

Unfortunately, the amount of evidence on this is still modest. Again, this is mainly due to the scarcity of data: not only matched employer–employee databases still do not abound, but also the information included in the available datasets often does not allow for a systematic discussion of the links between worker turnover and the dynamic features of firms and industries. The implication is that the available evidence related to such links was produced on the basis of very few data sources.

Keeping these limitations in mind, the following list summarizes a number of interesting results that can be drawn from the scattered empirical literature dealing with the role of firms and industries in explaining the inter-firm mobility of workers.

---

4In an early effort to distinguish between demand- and supply-side determinants of worker flows which did not take advantage of a matched employer–employee database, Davis and Haltiwanger (1992) were able to estimate approximately the impact of demand-led disturbances on worker mobility by combining plant-level data with information from different sources on the mobility in the labor markets. They suggested that 35–56% of the transitions between employment states were due to employment opportunities related to job creation and destruction. Although it is only an indirect estimation, this interval is actually not incompatible with the more precise estimations obtained on the basis of matched employer–employee data.

5Churning flows are typically computed at each period as the difference between total worker turnover (i.e., the sum of hires and separations occurring in that period) and the absolute value of net job changes. That is, \( CF = WF - |H - S| \), where \( CF \) are the churning flows, \( WF \) are the total worker flows \( (WF = H + S) \), \( H \) are the hires, and \( S \) stand for the separations in the period.
(i) **Firm size is not on itself a determinant of worker turnover:** some empirical studies have found a significant relation between these two variables: for example, using data for the USA during the 1980s, DiPrete (1993) regressed the number of worker separations per organization on the dimension of firms (and other variables related to industry and the worker type) and found an inverse relationship between the two variables; Tattara and Valentini (2005) also found that churning declines with the size of the employer. However, other studies seem to contradict those results: Burgess *et al.* (2001) found that churning is unrelated with size; in the same vein, using data for the UK in 1990–1991, Martin (2003) shows that the effect of firm’s dimension on worker turnover is not significant. While the comparability of this studies is somewhat hampered by differences in the variables they purpose to explain (separations, churning, total worker turnover) and in the data being employed (time and geographical scope, frequency of observations, unit of analysis, etc.), there are reasons to believe that firm size exert its influence on workers mobility by means of other organizational features—and therefore the statistical significance of their impact tends to wither as the relevant variables are included in the regressions.6

(ii) **Churning is positively related to firm’s growth:** the study by Burgess *et al.* (2000a) concludes that increases in firms’ employment lead to higher churning rates, while reductions in employment have the opposite effect; in explaining this result, the authors suggest that the expansion of firms’ workforce lead to an increase in bad matches, thus justifying the simultaneity of hires and separations at the firm level (again, the contraction of the workforce has the opposite effect).7 An increase in churning associated with expansion was also found by Tattara and Valentini (2005).

(iii) **Churning rates are higher for younger firms:** Lane *et al.* (1996), using data drawn from the quarterly employment and earnings records of the US State of Maryland, found that churning is slightly decreasing in the age of the firm.

---

6For example, in the study just mentioned, Martin (2003) finds that worker turnover is negatively related to wages, while not significantly related to firm size; on the other hand, wages have been shown to be systematically related with the size of firms (Oi and Idson, 1999); therefore, it may happen that a negative statistical relation between worker turnover and firm size will be rendered insignificant when wage is included as an explanatory variable in the regression. Of course this will depend on the factors that underlie the size-wage positive relation. Oi and Idson (1999) discuss alternative theories that account for such robust result.

7A few studies have tried to analyze the relative incidence of hires and separations as firms adjust to their new dimensions, but the results seem contradictory: Burgess *et al.* (2001) found that growing firms mostly increase their hiring and do not act to reduce turnover; declining firms generally maintain hiring but increase separations; on the contrary, Abowd *et al.* (1999), found that employment adjustments are primarily made by adjusting entry, rather than exit rates; the later result is consistent with the findings of Anderson and Meyer (1994) and of Albaek and Sørensen (1998). Tattara and Valentini (2005) suggest that the adjustment mechanisms are influenced by the structural evolution of the labor market.
Tattara and Valentini (2005) found that churning flows are higher during the first two years of activity. In order to analyze churning rates over the lifecycle of firms, Burgess et al. (2000b) divided firms into four lifecycle categories: (i) firms that survive 12 quarters or less; (ii) continuing firms within 12 quarters; (iii) noninfant firms within 12 quarters of death; and (iv) the rest (i.e., continuing mature firms). The results showed that the worker flow rate (i.e., churning rate plus job reallocation rate) decreases monotonically from category 1 (65.4%) to 4 (30.4%); but they also showed that hires and separations are important in every category, making churning flows more common across categories than job reallocation. This implies, for instance, that in continuing mature firms, which typically experience smaller changes in net employment (see Section 2.1), the weight of churning as a percentage of worker flows is particularly high. Still, the analysis of a specific cohort of firms confirmed the idea that churning rates tend to decrease with firm age.9

(iv) **Churning rates are a persistent, distinctive feature of firms:** in order to explain the heterogeneity among firms in the rates of labor market flows, Burgess et al. (2000a) ran separate regressions for job flows and for churning flows, including as independent variables time dummies, seasonal dummies, industry dummies, and employer dummies (fixed effects). While these regressors were able to explain only a very small part of the heterogeneity in job flows, about 50% of the variation in churning rates was explained by those variables, with employers’ fixed effects assuming particular relevance. These results clearly suggest that it is possible to identify firms that have systematically high churning rates and others which have systematically low churning rates (while the same cannot be said about the changes in net employment). In the same vein, Lane et al. (1996) have found that churning rates are positively dependent on past churning, which also point toward the presence of persistent differences between firms in relation to churning rates.

(v) **The incidence of churning is particularly high in some industries:** several studies have also revealed the presence of some industry specificities in churning rates. For example, Burgess et al. (2001) included industry dummies as regressors in an empirical model of the determinants of churning; their results show that, while the impact of most industries does not quantitatively differ in a significant manner, for a few other industries—namely, finance, insurance and real estate, and professional services—the effect on churning rates is particularly high.

---

8The same firm could be classified into different categories, depending on its current condition in different periods.

9There are two possible explanations for this: (i) the churning rate is reduced as firms ages due to better job matches, and (ii) high-churning firms have lower survival probabilities, so those firms that survive have typically low churning rates. Two pieces of evidence that will be discussed below—the negative relation between churning rates and firm survival, and the persistent heterogeneity of churning rates between firms—seem to favor the second explanation.
(vi) Industry turbulence seems to lead to higher churning (and not just to the reallocation of jobs): while the literature on job creation and job destruction clearly shows the impact of changes in industry structures in the reallocation of jobs (see Section 2.1), the evidence on the impact of industry turbulence on worker flows in excess of job reallocation is still rather scarce and unsystematic. For example, in their study of the California savings and loans industry, Haveman and Cohen (1994) have shown that organizational founding, organizational dissolution, and mergers and acquisitions have all had a significant impact on the mobility of executive employees between firms.

The six results listed above were drawn from a modest number of studies, with a limited time and geographical coverage, using often different techniques and measures, and should therefore be taken as essentially tentative propositions on the ways through which the evolution of firms and industries impacts on the mobility of workers across firms. Notwithstanding, they should help to design further empirical and theoretical investigations in this domain.

3. Evidence concerning the impact of worker turnover on the evolution of firms and industries

The relatively scarce availability of empirical work on the labor market determinants of industry dynamics tends to mirror the situation found above for the inverse direction of causality. Still, it is possible to find a few studies that have produced some evidence on the topic. Such studies can be divided in three groups: the first deals with econometric analyses of the impact of turnover on the performance of firms (in terms of growth, productivity, profitability, etc.); the second one consists on statistical analyses of firm survival, which include variables of workers mobility among the regressors; the third group corresponds to case-studies of specific industries.

Using a matched employer–employee database for France in the period 1976–1995, Kramarz and Roux (1999) estimated the effects of employee turnover on firm performance on the basis of firms’ tenure structure. They found that low turnover rates are associated with higher productivity, but a high turnover rate slightly favors profitability (suggesting the simultaneous presence of costs and benefits of labor mobility for firms). The simultaneity of harmful and beneficial turnover has also been identified by Garino and Martin (2005), using cross-sectional data for the UK; they found that the impact of worker mobility on firms would depend not only on the costs of recruitment and training, but also on the way wages are fixed in each context—turnover tends to have positive consequences for firms when wages are fixed exogenously.10

10The authors explain this on the basis of the idea that when firms are free to fix the wages, they minimize labor costs, leading turnover rates to increase over the optimal level.
Still in relation to the first group of empirical studies mentioned above, in analyzing the relation between churning flows (that is, worker turnover in excess of absolute net job changes) and job flows, Burgess et al. (2000a) have found that increases in churning flows typically lead to reductions in the size of firms. Using revenues, instead of total employment, as a proxy of firms’ size, Baron et al. (2001) also found that turnover (here understood as the proportion of individuals leaving the firm) has a negative impact on growth.

Among the second type of studies—those dealing with the relation between worker mobility and firm survival—Lane et al. (1996) have used a hazard rate model in order to test the prediction that high turnover firms will have lower survival rates. Together with other variables, the authors included as regressors the lagged churning rate of the firm (with various lags) to capture the effects of persistent churning. Their results strongly support the view that firms with high churning rates are less likely to survive than firms with lower churning rates, with the coefficients on churning rates lagged as much as three quarters showing to be consistently negative and significant. Burgess et al. (2000a) used the same data in order to analyze the relation between past churning and the probability of survival. They modeled the probability of survival at time $t$ as a function of the average churning up to $t$, and found that the impact of past average churning on firms’ survival was lower than current churning, but it was still significant. This result suggests that turnover is not simply the anticipation by workers of the future misfortunes of firms, and reinforces the argument that high churning may not be optimal for firms. Similar results on the impact of labor mobility on firms’ hazards were obtained by Mamede (2006) for the Portuguese knowledge-intensive business firms, who also found that the survival chances of firms would also depend on the educational level of those individuals entering and leaving the firms.

Evidence on less obvious impacts of labor mobility on the hazard rates of firms is provided by studies focusing on the relation between the probability of survival and the previous experience of firms’ founders. For example, Eriksson and Kuhn (2006) analyzed whether spin-offs11 take advantage of intangible assets, such as industry-specific knowledge, personal networks, or trust among its founders, in terms of their survival prospects (in comparison to other start-ups); they found that spin-offs were in fact associated with lower death risks than other types of entry. Pointing toward similar results, the literature on entry by spin-offs (see Klepper and Sleeper, 2005, and Helfat and Lieberman, 2002, for two influential papers related to the topic) has accumulated evidence on the relevance of the movement of workers out of incumbent firms and into new ventures in determining the evolution of industry structures.

11Spin-offs are understood in this context as new firms originating from within an existing company.
Finally, the centrality of labor market dynamics was shown to be a distinctive feature of a number of competitive contexts on the basis of industry-specific case studies. For example, the performance of firms was found to be very much affected by the capacity of firms to recruit skilled workers and to avoid poaching by competitors in industries such as professional services (Gallouj and Gallouj, 1996; Mamede, 2002) and hi-tech industries (Baron, 2004). In such cases, the patterns of worker turnover and firms’ persistently heterogeneous competences in managing human resources are clearly influential features in determining the patterns of industry evolution.

4. Theorizing on the coupled dynamics of industry structures and worker mobility

The development of research on industrial dynamics and on labor mobility has many interesting parallels. Empirical research in both fields has led to the identification of a number of (more or less robust) empirical regularities, on the basis of which existing theoretical models were tested and new models were developed, aiming at a better explanation of the regularities found in the data.

The previous sections have presented some pieces of evidence which suggest that the evolution of firms and industries and the mobility of workers might not be entirely independent phenomena. However, notwithstanding all the possible interdependencies between the two types of dynamics, there is a lack of theoretical discussion about the details of such coevolution and its implications. The present section identifies these gaps in the literature and discusses a number of mechanisms which can be at work in the context of such coupled dynamics, pointing toward alternative avenues for theoretical development.

4.1 Theorizing on the role of industry dynamics as a determinant of the inter-firm mobility of workers

Drawing on an extensive review of empirical studies concerned with the analysis of the stability and mobility of employment relations, Farber (1999) identified three central facts describing inter-firm worker mobility in modern labor markets: (i) long-term employment relationships are common (i.e., a significant proportion of workers are involved in durable employment relations); (ii) most new jobs end early; and (iii) the probability of a job ending declines with time (the relation is not necessarily monotonic—some studies find that the probability of a job change may first increase with tenure, before starting to decrease).12

12To provide an instance of such regularities, drawing on data from unemployment insurance systems in the USA, Anderson and Meyer (1994) have found that: most turnover (55%) is due to a minority of individuals (22%) who frequently change jobs; about 40% of employment relations last...
Different types of models have been put forward which can account for such statistical regularities. However, the most influential of such models have one thing in common: they typically abstract from the effects of industry turbulence on labor mobility (putting the burden of the explanation on factors belonging to the supply-side of the labor market, or on the quality of the match between employer and worker). Two paradigmatic models illustrate the point.

The first case consists in a model of worker mobility which is based on individual heterogeneity. The idea underlying this type of approach is that individuals have different propensities toward work and mobility (which may be captured, at least partially, in the empirical work by such variables as age, gender, ethnicity, education background, etc.). Farber (1999) presents a simple model of this type to show how individual heterogeneity can lead to the three regularities mentioned above: suppose there are two types of workers, which only differentiate by their turnover probability; in order to have a high percentage of long tenures, one only needs to assume that less turnover-prone individuals are highly represented in the population; since highly mobile workers have a lower probability of experiencing long tenures, most of the workers involved in durable employment relationships will be of the low-turnover type and, consequently, the probability of separations decreases with tenure; finally, since the less mobile workers are typically involved in long tenures, most new job vacancies will be filled by high-turnover individuals and, therefore, many new jobs matches will end early.

Jovanovic's (1979) model constitutes an alternative to the individual heterogeneity-type of explanations for the typical patterns of job turnover mentioned above. The building block of this model is the idea that the productivity of each particular job match is not known in advance—it is gradually revealed, since output constitutes a noisy signal of match quality. As the expectations of both firms and workers are updated on the basis of each period's output, both sides can decide whether to continue or to stop the employment relationship. Jovanovic's model is particularly successful in replicating the statistical regularities listed above, since it allows for a nonmonotonic relation between tenure and probability of turnover: initially, even if the observable output signals a bad-quality match, workers tend to remain in the firm since they know the signal is noisy; as time goes by, the assessment of match quality becomes more precise, leading either to a separation (because the match quality is too low) or to a permanent match (because its quality is high); thus, in an early phase more and more workers will decide to move, but on the other hand an increasing number of workers is involved in enduring employment relationships.

Both these approaches illustrate the tendency to abstract from demand-side disturbances as determinants of employee turnover. Such tendency is largely at odds

no more than 1 year; and the impact of tenure on the dissolution of job matches is negative (after controlling for other relevant factors, such as wage levels).
with the evidence discussed in Section 2, showing the extent to which labor market
dynamics are influenced by the turbulence experienced on the employers’ side.
Quoting Davis and Haltiwanger (1999, p. 2715), “It is now apparent, as perhaps it
was not a decade ago, that a satisfactory account of worker mobility dynamics in
market economies requires a major role for demand-side disturbances as well as for
supply-side and match-quality effects.”

In fact, the empirical studies discussed in Section 2 suggest that the structural
dynamics of industries can have a relevant impact on the mobility of workers in the
job market, not only directly but also in indirect ways. The observable patterns of
worker mobility seem to emerge from the consistent behavior of both workers and
firms, who systematically take into account the dynamic features of industries.
Moreover, it has been shown that varying degrees of worker turnover seem to be
a persistent characteristic of firms. Thus, the influence of persistently heterogeneous
employers acting in the context of changing industry structures emerges from this
discussion as central features in the understanding of the patterns of worker
turnover.

One can make sense of that interdependency by acknowledging that, at each
moment in time, the level of worker turnover may differ across industries or across
firms for such reasons as the following: high hiring and training costs may reduce
turnover, and such costs can vary across industries (or between different phases in
the lifecycle of the same industry) due to differences in the production process and
the nature of worker skill requirements; implicit contracts may develop between
firms and workers, and such contracts may vary with firm size and product demand;
there may also be economies of scale in screening devices, which may help increase
the initial quality of job matches in bigger firms; moreover, firms may learn to
develop more effective screening devices as they age, which means that turnover is
expected to be lower for older firms (Lane et al., 1996). All these examples suggest
that the type and amount of turnover expected to be found (and/or to be desirable)

It is worth noting that, following the empirical work that revealed the significance of “job creation
and job destruction”, many models have considered the interactions between the demand side of the
labor market and gross labor market flows (e.g., see the survey by Pissarides and Mortensen, 1999).
Those models, however, typically aim at explaining certain aggregate regularities, such as the
positive relation between wage and labor productivity, or the aggregate behavior of unemployment
and gross job flows—and not the central facts on the patterns of inter-firm worker mobility, which
were emphasized by Farber (1999), and which were presented before as the focus of the present
discussion. On the other hand, one can find models that focus specifically on worker mobility, while
considering at the same time the role of demand-side factors—as the one by Jovanovic and Moffit
(1990); this model nests match quality and sectoral shocks as determinants of labor mobility;
however, as is usually the case with models of this kind, it only considers productivity shocks that
are common to all firms in each sector; that is, although they take into account changes on the
firms’ side, such models still abstract from the role of industry turbulence in determining worker
mobility—which is the topic of interest here.

13It is worth noting that, following the empirical work that revealed the significance of “job creation
and job destruction”, many models have considered the interactions between the demand side of the
labor market and gross labor market flows (e.g., see the survey by Pissarides and Mortensen, 1999).
Those models, however, typically aim at explaining certain aggregate regularities, such as the
positive relation between wage and labor productivity, or the aggregate behavior of unemployment
and gross job flows—and not the central facts on the patterns of inter-firm worker mobility, which
were emphasized by Farber (1999), and which were presented before as the focus of the present
discussion. On the other hand, one can find models that focus specifically on worker mobility, while
considering at the same time the role of demand-side factors—as the one by Jovanovic and Moffit
(1990); this model nests match quality and sectoral shocks as determinants of labor mobility;
however, as is usually the case with models of this kind, it only considers productivity shocks that
are common to all firms in each sector; that is, although they take into account changes on the
firms’ side, such models still abstract from the role of industry turbulence in determining worker
mobility—which is the topic of interest here.
in different firms, in different industries, or in different phases of the lifecycle of firms and industries may vary in some consistent manner.

4.2 Theorizing on the role of worker mobility as a determinant of the evolution of firms and industries

As in the case of labor mobility, the empirical evidence drawn from several studies on industry dynamics has allowed the identification of some statistical regularities (for surveys see Geroski, 1995; Dosi et al., 1997; Caves, 1998), including the following: the entry and exit of firms are two frequent, and very often correlated, phenomena; the distribution of the size of the firms is typically biased toward smaller scales; new firms are smaller than the average incumbent, have a small probability of survival, and those that survive grow faster than the average; the variability in firms’ growth rates diminishes with size; several industries experience shake-outs in the number of firms, after reaching a peak in the number of incumbents.

Many formal models have integrated those (and other) regularities in their assumptions and/or replicated them in their outcomes, providing alternative explanations for the observable patterns of industry dynamics. In spite of the diversity of the causal mechanisms put forward in those theoretical exercises, industry dynamics models tend to focus on technological or financial determinants of changes in the structure of industries, abstracting from the possible impact of labor market dynamics on those changes.

Two paradigmatic contributions in this domain illustrate the point. In the evolutionary models of Nelson and Winter (1982), the selection of firms is determined by their innovativeness, which is a stochastic function of firms’ investments in R&D; innovative behavior of firms hence determines the structure of the industry and its evolution, which follows the above mentioned patterns for the relevant part of the space of parameters. In contrast with these models, in which agents actively invest in learning, Jovanovic’s (1982) model is able to replicate many of the statistical regularities of industry dynamics assuming instead that firms are born with a certain level of efficiency; such level is not known with certainty by firms when they enter the market; over time, production outcomes gradually reveal the true efficiency levels, leading to decisions by firms to either expand or contract (and eventually exit the market). The same type of emphasis on information updating or on technological learning—and the absence of labor mobility among the causal mechanisms—can be found in other reference models of industry dynamics, including the ones by Hopenhayn (1992), Jovanovic and McDonald (1994), Ericson and Pakes (1995), Klepper (1996), and Winter et al. (2003).

Although such models were relatively successful in replicating a number of statistical regularities associated with industry dynamics, they shed little or no light on the ways through which changes in industry structures may be influenced by the
mobility of individuals in the labor markets. And, still, there are reasons to suspect that a relevant part of the picture is thus being left aside.

As discussed in Section 3, the few studies available which tried to assess the impact of labor market dynamics on industry evolution point toward the simultaneous presence of both positive and negative effects of worker turnover on firm’s performance. Any attempt to incorporate such impacts on a theory of industry evolution must start from identifying the specific mechanisms through which those impacts are produced.

A detailed discussion on how worker turnover can induce both costs and benefits to organizations is found in Staw (1980), providing some rationale for the (admittedly scarce) evidence discussed above. Turnover-related costs include: costs of selection, recruitment, and training (which are expected to be especially high for complex jobs in the context of tight labor markets, in particular for firms which cannot rely on dedicated departments and/or internal mobility); operational disruption (particularly when turnover affects central functions in the context of highly interdependent structures); demoralization of organizational members (when turnover affects group cohesion). While organizational costs of worker mobility are often emphasized, turnover may also be beneficial to the performance of organizations in several ways, such as: new hires can be associated with more motivated, more competent, and more educated workers; the exit of workers (in the form of either fires or quits) is one of the possible solutions to entrenched organizational conflicts; worker turnover (both inwards and outwards) can lead to a diversification of the external links of organizations, with benefits in terms of access to different types of resources.

In this context, it is worth noting that the idea of worker turnover having deleterious consequences to (and which are somewhat anticipated by) firms has provided the basis for the explanation of some well-known phenomena in labor economics. In particular, efficiency wage theories (Akerloff and Yellen, 1986) incorporate the assumption that employee turnover is reduced by increasing current and (expected) future wages and other benefits. This explains why wages are often higher than expected, or why incentive regimes are particularly generous in rewarding tenure (as found, for example, by Møen, 2005, in the case of technical staff in R&D-intensive firms, where the wage-tenure profile is particularly steep), a situation that tends to occur whenever reducing turnover rates is highly beneficial to the firm (e.g., by promoting investments in firm-specific human capital, and/or reducing the costs of searching and recruitment).

The fact that firms respond to the risks posed by employee turnover resorting to internal incentive schemes may suggest that, ultimately, this renders the mobility of workers irrelevant (since firms would be always optimally responding to the possibility of turnover, and would therefore have the amount of worker turnover which was considered the most appropriate). However, the fact that firms display persistent differences in their propensity for labor mobility (see Section 2.2) may be
an indication that the later is not always the result of optimal turnover strategies, but rather an essentially idiosyncratic, organizational trait—and, therefore, labor mobility may indeed autonomously contribute to the dynamics of industry structures.

4.3 Possible alternative strategies for theoretical developments

Examples of industries where competition is based on recruitment (to borrow the expression used by Sørensen, 2004) were given above. Those examples suggest that, at least in such contexts, the evolution of industry structures can be highly influenced by the patterns of worker turnover and by the heterogeneous competences of firms in this domain. On the other hand, the highly turbulent character of some of those industries (especially those in the early phases of their life-cycles), implies that the movement of workers between firms is strongly influenced by the dynamics of the relevant population of employing organizations. That is, one can expect to observe causality running in both ways, suggesting the opportunity for—and the usefulness of—an integrated approach to the dynamics of industry structures and labor mobility.

Nevertheless, as could be expected from the discussion above, examples of theoretical analyses taking into account the joint dynamics of industry structures and labor mobility do not abound in the literature. Given this scenario, I discuss below two possible strategies for the development of theoretical approaches to the coupled dynamics of industry structures and labor mobility.

The question of interest here is: what are the conceivable mechanisms through which industry turbulence (understood as the entry and exit of firms, changes in sizes and in market shares, changes in property control, evolution in industry concentration) and labor market mobility would mutually influence each other? Answering this question is equivalent to sketch the main features of alternative integrated models dealing with the problem at hand.

One possible strategy for the development of such approach to this problem is suggested by bridging the literature on organizational ecology (or corporate demography—for an extensive review see Carrol and Hannan, 2000) and the research on internal organizational demography (Pfeffer, 1985). An example of such bridging can be found in Haveman (1995), who starts from the idea that the founding, dissolution, and merging of organizations has systematic impacts on firms’ internal demographic composition (namely, in terms of tenure distribution). She suggests, for example, that while short-tenured employees are more likely to exit firms in general (because they may not fit their jobs or firms well, or because they

14Burgess et al. (2000a) is the only paper I am aware of that tried to assess empirically such type of two-way causality. Using vector autoregression analysis, the authors show that the relation between job flows (i.e., absolute net changes in total employment) and churning flows (i.e., worker flows in excess of job flows) is in fact bidirectional.
have developed little firm-specific capital), moving into new ventures is an especially attractive opportunity for long-tenured employees because they possess the reputations, expertise and external contacts on which new ventures rely. Thus, in periods of high entry rates the proportion of long- to short-tenured employees leaving established companies will increase (and since not all long-tenure individuals leave their firms, the tenure dispersion in organizations will increase).

Although Haveman’s paper does not discuss the feedback effects from changes in organizational tenure distributions to the evolution of firms and industries, this has been a central concern for research done in the field of internal organizational demography. According to Pfeffer (1985), two central suggestions have been put forward (and empirically investigated) by organizational demographers: (i) the idea that tenure (and other demographic characteristics) strongly influence the managerial competences of individuals; and (ii) that the distribution of the competences among management teams has a significant impact on the performance of firms (the impact in terms of relative performance may depend on the distribution of competences in the management teams of the direct competitors, as suggested by Sørensen, 1999b).15

In sum, on the one hand, as Haveman (1995) suggests, different dimensions of industry turbulence (entry, exit, merger, etc.) lead to selective changes in the internal demography of firms (namely, in terms of tenure distributions); on the other hand, as organizational demographers emphasize, such changes imply a reconfiguration of the set of competences in firms, which may be expected to affect the performance of firms (and, when the population of firms in an industry is jointly considered, to affect the structure of the industry as well). Thus, the coupled dynamics would be here obtained on the basis of the causal sequence “specific elements of industry turbulence—selective worker turnover—changes in internal demography of firms—differential impacts on the performance of firms—further industry turbulence—...”

Another possible bidirectional link between industry dynamics and worker mobility is related to the role of social networks as part of the structure of both the industry and the labor market. The fact that social networks can, and often do, influence the dynamics of labor markets has long been emphasized by economic sociologists (Granovetter, 1995), and has been increasingly discussed by labor economists (e.g., Montgomery, 1991; Bentolila et al., 2004; Pellizzari, 2004). Studies within this tradition have revealed that employers and employees tend to know (or, at least, have information about) each other even before the beginning of their labor

15The idea that the individual characteristics of top managers can have strong implications for the strategy and performance of firms has also been central to the research tradition in management studies frequently referred to as research on the “upper echelons” (for a seminal paper, see Hambrick and Mason, 1984).
relation; and that social networks are extensively used by both firms and workers to find jobs and fill vacancies.16

Granovetter’s approach has a clear dynamic flavor: worker mobility is not only (partly) determined by the social structure, but also it helps in turn to change the social structure itself—since new personal links are being established as workers move between firms. And while his framework does not take into account the dynamics of firms and industries, it is not difficult to think of ways in which social networks, industry structures, and worker mobility can actually coevolve. For example, the mobility of workers (the crucial element in Granovetter’s analysis) often has important informational consequences to firms—e.g., firms have an incentive to hire individuals who are known to their employees (since this would facilitate the access to detailed information about the competences and personality of prospective employees); and since social links are expected to be more easily established between individuals with similar demographic characteristics (age or tenure cohorts, educational or socio-economic background, etc.), one can expect that the impact of social networks on the mobility patterns will also affect firms’ performance and, therefore, help to shape the evolution of industries; to the extent that social links can be formed when individuals work for the same firms, the evolution of industries feeds-back on the network structure and on the patterns of future mobility. This causal sequence, in which “structure shapes mobility, mobility shapes structure” (where “structure” refers both to social networks and populations of firms), may apply not only to the informational role of networks in determining job matches, but also to workers’ preferences concerning their choices of employers (e.g., all else being equal, individuals will prefer to work for those organizations in which they have acquaintances).

The two types of causal sequences presented above are only instances of alternative processes that may underlie the coevolution of industry structures and labor mobility. Whether these specific processes will show to be relevant in jointly explaining the patterns of industry dynamics and of worker mobility observed in specific industrial contexts, is left to investigation. In any case, those examples may help to stimulate the further development of models dealing with the coupled dynamics under discussion.

16Granovetter (1995) tends to emphasize the benefits of social networks for individual (not necessarily social) outcomes in the labor market; namely, he suggests that: information given by personal acquaintances about the nature of a job is often considered more reliable; friends may facilitate individual integration and learning in organizations; having personal acquaintances among colleagues can facilitate the access to promotion and other discretionary benefits (especially, if those acquaintances are well positioned in the organizational power structure, and if contracts are more difficult to be drawn exhaustively and enforced). On the other hand, it has been shown that in some contexts jobs found through acquaintances may be actually associated with lower wages—see, e.g., Bentolila et al. (2004); Pellizzari (2004).
5. Conclusions and opportunities for future research

The idea pervading this article is that the mobility of workers in the labor markets and the patterns of industry evolution can, and often do, influence each other, and that both empirical and theoretical research in those two domains should take into account such interdependencies.

The contributions of this article were presented in five steps: first, the scattered evidence pointing toward the impact (both direct and indirect) of industry dynamics on labor mobility was identified; second, evidence supporting the inverse direction of causality—i.e., the influence of worker turnover on the evolution of firms and industries—was gathered; third, by referring to paradigmatic models of both industry dynamics and labor mobility, it was shown that the theoretical literature in these two research fields have essentially ignored the possible interdependencies between the two types of dynamics; fourth, the article discussed a number of causal mechanisms which can give rise to such interdependencies; finally, some alternative possibilities for theoretical developments were put forward, by bringing together in the same framework part of those mechanisms.

One of the main messages arising from the discussion above is that opportunities abound for further empirical and theoretical work on the interplay between industry dynamics and labor mobility. The possibilities for empirical developments are growing as more and more matched employer–employee database are becoming available. While this type of data has been increasingly used within labor economics, its potential for the advance of knowledge in the field of industrial dynamics is still rather unexplored. And even in labor economics such data could be used to analyze more systematically the influence of industry turbulence on the patterns of worker mobility.

The following list of questions informs an agenda for future empirical research:

– Do workers cluster within specific firms according to their propensity to turnover? If yes, what are the dynamic features of such firms?
– Do highly mobile labor markets lead to lower survival chances for some types of firms? What are the main characteristics of those firms?

17One example of a fruitful use of matched employer–employee data in the field of industrial dynamics—which is not exactly related to the problem dealt with in this article—was given by Benedetto et al. (2004). Empirical work in industrial dynamics has often used administrative data to follow the firms’ trajectories trough time; one problem with such data, which has been recurrently identified (but not satisfactory solved), is the fact that entry and exit can be mistakenly measured, since simple changes of ownership or legal form of organizations may modify the administrative identifiers with no other change in economic activity. In that paper, the authors describe how that type of datasets can provide information about the flows of clusters of workers across business units in order to identify longitudinal linkage relationships in business data.
How are worker turnover and its determinants (such as the costs of hiring and training, the efficiency of screening devices, wages levels, among others) related with the size and age of firms, and what impact does it have on the postentry performance of firms?

Are entry rates determined by the patterns of worker mobility? More specifically, can one expect higher entry intensities in markets which are populated by workers with a higher tendency to move between employers?

What are the indirect impacts (i.e., beyond direct job creation and destruction) of entry, exit, expansion, and contraction of firms on worker mobility, namely in terms of vacancy chains?

Do rates of worker turnover systematically vary between different phases of industries’ life-cycles? If yes, what are the main determinants of such variance?

To what extent the impact of worker turnover on the performance of firms depends on the different phases of their lifecycle, and on the type of industry?

To what extent differences in the way firms adjust their sizes (which are highly influenced by national laws and regulations) have an impact on the dynamic patterns of both industries and labor markets?

Such questions remain largely unanswered, and interesting results could arise by investigating them empirically. The identification of some basic empirical regularities concerning the mutual influence between industry turbulence and worker mobility would help the theoretical effort in throwing light on the main mechanisms at work.

One may question the scope of applicability of such integrated models. To be sure, the movement of workers between firms may be only marginally relevant to the evolution of some industries, particularly those that rely on a low-skilled, homogeneous workforce. Similarly, the role of industry dynamics in determining the patterns of labor mobility is not expected to be high if the industry’s turbulence is minimal.

However, taking into consideration the mutual influences between changes in industry structures and the mobility of workers may be crucial to the understanding of the dynamic patterns observable in several contexts. In particular, an integrated approach to industry dynamics and labor mobility will be particularly appropriate for the analysis of industries in the early phases of their lifecycles (when structural turbulence tends to be highest) and in which competition is strongly based on the recruitment of highly skilled workers. These criteria would often include some of the most dynamic industries in the contemporary societies—such as higher education, biotechnology, consultancy, law firms, among others—to which the prevailing models of industry dynamics and labor mobility are not particularly well suited. At least in such contexts, if not in others, theoretical and empirical inquires of industry evolution which abstract from the role of labor market dynamics—or vice-versa—risk missing the main elements of the dynamic picture they propose to explain.
Acknowledgements

This version of the article has benefited from comments and suggestions made by Faiz Galouj, Franco Malerba, Helena Lopes, Joao Rodrigues, Michele Pellizzari, and two anonymous referees, to whom I am thankful. The usual disclaimer applies.

Funding

Portuguese Science and Technology Foundation (grant ref. SFRH/BD/8765/2002, partial).

Address for correspondence

Ricardo Mamede, ISCTE - Departamento de Economia, Av. Forcas Armadas, 1649-026 Lisboa, Portugal. e-mail: ricardo.mamede@iscte.pt

References


