Internal Labor Markets: A Worker Flow Approach*

Ingrid Huitfeldt† Andreas R. Kostøl‡ Jan Nimczik§ Andrea Weber¶

This version: June 15, 2020

Abstract: This paper provides a new method to study how workers' career and wage profiles are shaped by internal labor markets (ILM) and job hierarchies in firms. Our paper tackles the conceptual challenge of organizing jobs within firms into hierarchy levels by proposing a data-driven ranking method based on observed worker flows between occupations *within* firms. We apply our method to linked employer-employee data from Norway that records fine-grade occupational codes and tracks contract changes within firms. Our findings confirm existing evidence that is primarily based on case studies for single firms. We expand on this by documenting substantial heterogeneity in the structure and hierarchy of ILMs across a broad range of large firms. Our findings on wage and promotion dynamics in ILMs are consistent with models of careers in organizations.

Keywords: Internal Labor Markets; Organization of Labor; Wage Setting.
JEL codes: J31, J62, M5

1 Introduction

While a large empirical literature studies how worker mobility across firms and industries affects wages (e.g., Krueger & Summers, 1988 and Abowd *et al.*, 1999), there is much less evidence on how worker careers and wages are shaped by the internal labor markets of firms. Early theoretical work characterizes internal labor markets as collections of jobs in firms (see, e.g., Doeringer & Piore, 1971), where promotion along internal job ladders follows fixed rules set by the firm. These rules can be used to induce effort (Lazear & Rosen, 1981) and to allocate talent within the firm (e.g., Sattinger, 1975 and Gibbons & Waldman, 1999b). The lack of empirical evidence is unfortunate as life-cycle labor market outcomes of workers are shaped by factors from both internal labor markets and external labor markets (Topel & Ward, 1992).

One explanation for the scarcity of evidence is that classifying internal labor markets and job ladders within firms is conceptually and practically challenging. Except for management positions, the main purpose of occupational codes or job titles is to describe tasks and they are not designed to give a clear representation of the hierarchical order within the firm. As a result, the reconstruction of firm-specific job ladders

*This project received financial support from the Norwegian Research Council through grant 227115. We would like to thank our discussant Thomas Lemieux, as well as Patrick Kline and Daniel Silverman for very constructive feedback and participants in several seminars and workshops for many useful suggestions.

†Statistics Norway. E-mail: ingrid.huitfeldt@ssb.no
‡Arizona State University, W.P. Carey School of Business; Statistics Norway; IZA. E-mail: andreas.kostol@asu.edu
§ESMT Berlin; IZA. E-mail: jan.nimczik@esmt.org
¶Central European University; IZA. E-mail: WeberA@ceu.edu

requires a substantial amount of knowledge about the organizational charts that map the hierarchy structures in companies. The second, more practical problem concerns data availability. Most large-scale linked employer-employee registers do not record detailed occupations or occupational changes within firms. The tedious process of collecting data and evaluating job titles has limited the existing evidence to case studies from particular firms. While Baker *et al.* (1994) provided the first evidence on job ladders using changes in job titles within one particular firm, evidence from a broader population of firms is still missing.

In this paper, we contribute to the conceptual issue by proposing a two-step method of measuring the internal career structure of firms. Our method is based on observed worker flows between fine-grained occupational codes or job titles within firms. We combine recent advancements in panel data estimators and clustering techniques, which have been applied to worker transitions between firms, to estimate the organizational structure *within* firms. The key to our method is having access to detailed matched employer-employee data with information on changes in job titles of individual workers in firms. In the first step, we identify firm-specific networks of occupations based on observed occupational transitions within firms. Occupations that are connected by worker flows form internal labor markets (henceforth ILMs). We allow these networks to be segmented, so that a firm can consist of multiple ILMs.[1] Our method allows for measurement error in the coding of rare occupational transitions, which could potentially influence the shape of the ILMs. In particular, we apply a pruning algorithm related to the method used by Kline *et al.* (2019), which checks if removing a single worker breaks an ILM into further sub-markets.

In the second step, we establish a ranking of occupations within the ILM by exploiting the direction of internal network links and the flow frequencies between occupations. Our approach builds on an intuitive idea by Baker, Gibbs and Holmstrom (1994, henceforth BGH): If many employees move from occupation b to occupation a within a firm but few, if any, move from a to b, the relative flows indicate that occupation a ranks higher than b. A clear path from c to b to a indicates a strong hierarchy in the internal labor market. We call such pattern a job ladder. In contrast, an intransitive ranking with flows in multiple directions indicates flat organizations where job rotation plays a bigger role. To systematically estimate occupational ranks within each firm, we apply a ranking algorithm that computes a hierarchy score based on the fraction of upward moves along the job ladder. We then minimize the number of downward moves over all possible rankings following the Markov Chain Monte Carlo procedure in Clauset *et al.* (2015). Finally, we group occupations into hierarchy levels to distinguish lateral moves from promotions and demotions. We cluster occupations with similar estimated ranks into a hierarchy level using a k-means clustering algorithm, and follow the data-driven approach in Bonhomme *et al.* (2019) to choose the optimal number of levels. This allows us to summarize the organizational structure of firms by a single statistic, the number of hierarchy levels.

We overcome the measurement challenge by leveraging rich administrative data from Norway with information on worker mobility both across firms and between job titles (fine-grained 7-digit occupations) within firms. More than 50 percent of all job changes in our data occur within firms, and within-firm job changes are associated with larger wage increases than firm switches on average. This observation motivates our analysis of promotion dynamics in ILMs.

[1]The counterpart to an internal labor market is a connected set of firms based on between-firm worker mobility in the terminology of Abowd *et al.* (1999) and Abowd *et al.* (2002).

We implement our approach on a set of about 2,000 large private sector firms in Norway. Our empirical analysis can be summarized with three broad conclusions. First, we document that the vast majority of Norwegian firms have one large ILM and a few single occupations that are unconnected from the main ILM. This holds even as firm size increases. Larger firms have more single occupations, but overall the major part of workers and occupations are included in one ILM. The structure of the ILM differs widely across firms, however. While we can describe ILMs by up to 19 hierarchy levels, about half of our sample of firms have three or four levels of hierarchy. The length of the hierarchy increases in firm size and the number of occupations employed by the firm.

Second, the ILMs we identify in the data are broadly consistent with theories of internal labor markets (e.g., Doeringer & Piore, 1971) and of hierarchies in the organization of labor (e.g., Garicano & Rossi-Hansberg, 2006). According to these theories, workers enter the firm at the bottom of the hierarchy and then move up the career ladder towards more complex jobs over time. We find several pieces of evidence showing that our estimated job ladders are consistent with these models. Hierarchies have a pyramidal structure: The largest number of workers are employed at the bottom and employment declines toward the top. The share of management positions increases over the hierarchy, indicating that task complexity increases. In line with the idea of entry at the bottom, the share of hires from the external market is highest at the lowest hierarchy level, but it declines toward the top of the hierarchy where the majority of workers are hired internally. The idea of promotions is also reflected in the increasing average tenure over the hierarchy. These basic descriptives are further confirmed by tracking individual workers after entry into the firm. Within 6 years of entry, about a third of these workers experience a promotion to a higher level, while demotions are much rarer.

Third, we find evidence in support of a link between the internal hierarchy and the wage setting process via a system of promotions. We document that the average log wage increases almost linearly with the hierarchy level for the average firm and also within firms. Individual level regressions of log wages with firm by year fixed effects, show that hierarchy levels predict wages even after flexibly controlling for age, education and tenure. The finding that wages increase across hierarchy levels resembles the evidence by Baker *et al.* (1994) and is in line with tournament theory (e.g., Lazear & Rosen, 1981; Waldman, 1984) which predicts an increase in the wage spread between hierarchy levels towards the top of the job ladder. We also find that average person fixed effects from AKM wage regressions increase in hierarchy levels. This result is consistent with sorting due to incentives from promotions tournaments, where higher ability workers climb the hierarchy ladder faster. Alternatively, the finding is also consistent with assignment models where firms use the hierarchy to allocate more talented individuals towards higher levels.

Our paper is closely related to a literature that studies organizational structures and career progressions within firms. In the original theory of Doeringer & Piore (1971), internal labor markets are characterized by ports of entry and exit as the only points of interaction between internal and external markets. In contrast to a spot market where workers are paid their marginal productivity, workers have careers within firms and receive wages that are attached to the job characteristics and are not subject to influence from the outside market.[2] Baker *et al.* (1994) offer a first empirical assessment of these concepts based on personnel data from a specific firm. Our evidence is consistent with the concept of ILMs discussed in the literature, and we

[2] Gibbons & Waldman (1999b) extends the theory by providing an integrated theory of job assignment, human capital attainment and learning.

contribute to it in two ways. First, our data-driven method can be applied and used to identify the hierarchical structure of ILMs of any particular firm. Second, linked employer-employee data allows us to provide novel evidence on the hierarchical structure of ILMs across a variety of firms.

A small literature experiments with alternative measures of internal labor markets and the assessment of labor market impacts of ILM's beyond single firms. Early studies rely on survey reported promotions by either workers or employers (McCue, 1996), while others rely on broad occupational groups (Lazear & Oyer, 2004 and Van der Klaauw & Da Silva, 2011) or internal connections between larger organizational units such as business groups or establishments (e.g., Cestone *et al.*, 2019 and Giroud & Mueller, 2019). In contrast to these studies, our method provides a much more fine-grained measure of ILMs that goes down to the occupational level but also captures connections across organizational units.

Empirical evidence from case studies of individual firms has triggered a broad variety of theoretical approaches to explain the hierarchical structure of jobs and wages (see e.g., Lazear & Shaw, 2007). In the seminal tournament models of Lazear & Rosen (1981) and Waldman (1984), for example, firms introduce wage spreads between different hierarchy levels with the aim of eliciting worker effort, incentivizing human capital investment, or preventing the worker from being poached by competitors. However, the extent to which observed wage differentials reflect differences in ability or incentives has proven hard to disentangle empirically (for a survey, see Chiappori Pierre & Salanie, 2003).[3] While testing the competing theories is beyond the scope of this paper, our method develops empirical tools that permit a fresh take on longstanding questions, such as the assessment of the relationship between the firm's hierarchy structure and the internal structure of wages (e.g., Lazear & Shaw (2009)).

The paper proceeds as follows. Section 2 describes our data sources and the institutional background. Section 3 introduces our methodological approach to identifying internal labor markets and job ladders. Section 4 discusses the properties of estimated internal labor markets and job ladders in the sample of large Norwegian firms. Section 5 discusses the robustness of our estimation methods to alternative assumptions about measurement error in the data and alternative clustering approaches. Section 6 concludes.

2 Data and Institutional Background

This section describes the administrative matched employer-employee data that we use, explains our key variables, and provides institutional background on labor markets in Norway.

2.1 Data

We use administrative matched employer-employee data from Norway. The data can be linked by unique and anonymized identifiers for every labor force participant, firm, and establishment. The Norwegian employer-

[3] Consistent with predictions from tournament theory, where wage spreads must rise with the number of workers in order to compensate for the increased competition for higher-ranked jobs, Gabaix & Landier (2008) show using cross-country data that the size of firms explains the bulk of differences in compensation. Eriksson (1999), Bognanno (2001), Garicano & Hubbard (2007) and DeVaro & Kauhanen (2016) find empirical support for the theoretical predictions of these models. Bertrand & Schoar (2003) document important manager fixed effects in several dimensions of firms practices, and Lazear *et al.* (2015) use company-based data and estimate that CEOs are on average 1.75 times as productive as the average worker, suggesting that incentives explain the bulk of the variation in compensation in the particular firm.

employee register includes virtually all employment contracts from 2006 to 2014, except for contracts with fewer than 4 hours of work per week or below 10,000NOK per annum. The contracts are reported by the employer to the authorities at the end of the year.[4] The reported contracts include information on the dates of alterations to the terms of a contract, the corresponding wage, industry and occupational codes, geographic location of the workplace, and tenure. Moreover, the data allow us to construct time series of monthly earnings for each worker, and to track all cases where a worker switches her firm, establishment, or occupation.

2.1.1 Variables

From our data source, we extract all transitions of workers between occupational codes, both within and across firms. We keep track of worker earnings and other characteristics. This section describes our main variables.

Occupations. The key to our approach is fine-grained occupational codes that can be used to describe the variety of jobs or positions at a given firm in detail. We use 7-digit occupational codes based on the international standard classification of occupations by the international labor organization (ISCO). There are about 5,000 different occupations in the Norwegian version, where some job descriptions have been adjusted to meet Norwegian standards and occupational licensing rules. Given the detailed and fine classification of occupations, we use the terms occupation, job (title), or position interchangeably. For some jobs, the occupational descriptions include information about the rank of the occupation in the hierarchy, e.g. assistants, mid-level managers, top-level management, or members of the executive board. For the majority of jobs, however, it is not possible to classify the occupational code into a hierarchy structure based on the occupational description. In Section 3, we therefore introduce a data-driven method to classify occupations based on the flow of workers across 7-digit occupations within firms. These flows are a reliable measure of job changes since our data record every change of contract - including a change in the occupational code. Our data allows us to observe transitions within organizations that cross different establishments as well as across occupations, e.g. a person working as a systems engineer in one plant who becomes an operations manager at another plant, and moves on to a central position at the headquarter.

Wages. Our measure of wage is the natural logarithm of the average monthly earnings of a worker in a firm. In our empirical application, we primarily use log wage. We also use residualized log wages when estimating fixed individual and firm effects. We residualize wages by regressing the log wage on a flexible specification of calendar year indicators that capture common year effects and include individual characteristics, such as dummies for each year of schooling, whether a person is married, a dummy for each number of children below age 18, gender, and each age category.

[4] We exclude data prior to 2006 as the occupational codes were incomplete in some firms. From 2006 to 2014 earnings has reported in spells, where every change in the contract was recorded. The process of reporting income changed in 2015, and has since January 2015 been based on complete monthly payments. We exclude observations after 2014 to avoid mechanical changes to the wage structure and to handling missing contracts prior to 2015.

Worker and Firm Heterogeneity. Following the seminal work of Abowd *et al.* (1999, henceforth AKM), we decompose wages into additive fixed effects that represent unobserved worker and employer heterogeneity. Let w_{ijt} denote the log residualized wage of individual i in year t and in firm j. The AKM wage model with firm fixed effects is described by

$$w_{it} = \alpha_i + \phi_{j(i,t)} + r_{it}, \tag{1}$$

where α_i is a time-invariant person effect for worker i, and $\phi_{j(i,t)}$ is the permanent firm fixed effect of firm j that employs i in year t. The time-varying residual component r_{it} is assumed to be uncorrelated with worker mobility. We estimate this model on our full sample and use worker and firm effects as additional characteristics that describe unobserved heterogeneity.

Demographics. To capture complete information on workers' geographic locations and other socio-economic variables we link the matched employer-employee data with longitudinal administrative registers provided by Statistics Norway. These administrative data sources cover every Norwegian resident from 1967 to 2014 and contain individual demographic information like sex, age, zip codes, and education.

2.1.2 Sample Selection

Our empirical analysis focuses on private sector firms. A firm can consist of multiple establishments (or plants). Since our main interest is in the organization of jobs in firms that offer career possibilities to workers, we require a firm to have at least 30 employees at some point during the period from 2006-2014 and at least 10 internal movers over the sample period. We further restrict our sample to firms with at least 15 external hires over the period to reduce bias that arises from limited mobility in the estimation of the AKM wage model.[5] This leaves us with a sample of slightly more than 2000 firms that employ about 100 workers on average.

The sample of workers includes every full-time employed male and female aged 20 to 61. This restriction is customary in the literature and avoids issues related to work hours and labor force participation. If a worker has multiple employers over the year, we choose the employer in the last month observed, and if the person has multiple employers in the last month, we select the employer with the highest total earnings in a particular year. We also exclude transitions of workers who have a period of more than 6 months of non-employment between two positions or firms.

2.1.3 Descriptive Statistics

Table 1 provides some descriptive statistics on the workers in our data. We split the sample into stayers and movers. Among movers, we also distinguish between those who switch between firms and those who switch their occupation within firms. The majority of workers are not moving from one year to the other. While a large share of the literature in labor economics focuses on external job transitions, our data show that internal transitions are important: Among movers, there are more internal than external moves (53 percent versus 47

[5] See e.g. Andrews *et al.* (2008), Kline *et al.* (2019), and Lamadon *et al.* (2019) and the discussions therein.

Table 1: Descriptive Statistics: Movers vs. Stayers

	A. Stayers		Movers: B. Internal		Movers: C. External	
	mean	st.dev	mean	st.dev	mean	st.dev
Worker outcome:						
Log wage	10.56	(0.501)	10.74	(0.483)	10.46	(0.507)
Wage growth	0.0677	(0.231)	0.0830	(0.258)	0.0684	(0.447)
Share positive growth	0.607		0.749		0.575	
Share negative growth	0.393		0.251		0.425	
Mean person fixed effect	9.840	(0.324)	9.925	(0.305)	9.809	(0.294)
Worker characteristics:						
Age	39.99	(11.2)	40.42	(10.7)	36.56	(10.2)
Female	0.297		0.304		0.282	
Married	0.520		0.553		0.424	
6-9 years education	0.0471		0.0390		0.0474	
10-13 years education	0.620		0.563		0.594	
14-16 years education	0.333		0.398		0.359	
Firm Characteristics:						
Total Employees	1649.6	(3513.6)	4029.7	(6185.7)	1439.7	(3061.4)
Number of Job Titles	10.85	(12.75)	9.979	(12.49)	10.16	(12.18)
Mean firm fixed effect	0.609	(0.144)	0.655	(0.145)	0.600	(0.175)
Person-years	3,296,544		264,301		238,587	

Notes: This table documents the characteristics of movers and stayers. The sample is restricted to private sector firms with at least 30 employees and at least 10 job transitions within the firm during the period 2006-2014. We further restrict our sample to firms with at least 15 external hires over the period. The sample of workers includes every full-time employed male and female aged 20 to 61. If a person has multiple employers over the year, the data includes the employer in the last month observed; and, if the person has multiple employers in the last month, we select the employer with the highest total earnings in a particular year. We also exclude workers who have a period of more than 6 months in non-employment between an external and internal move. Wage is at nominal levels. Observations are worker-year.

percent). Moreover, internal movers experience a higher average wage growth than those who do not move and those that stay. Internal transitions are associated with a much lower share of negative wage changes. Internal movers are positively selected as indicated by a higher average person fixed effect in our AKM wage decomposition.

Consistent with the previous literature, we observe that workers who move across firms are younger (see e.g., Neal, 1999). We also observe that those who move internally are more educated, and are moving within firms that employ many workers. Firms with internal movers are characterized by higher AKM firm effects compared to those of stayers and external movers. These facts motivate the remainder of the paper, where we try to open the black box of the firm and aim to measure and understand internal labor markets.

2.2 The Norwegian Labor Market

The Norwegian labor market is characterized by a combination of institutional regulation and flexibility. Hiring and firing practices follow European labor law. Firms can hire employees on either fixed-term or permanent contracts, where a permanent contract typically entails a probationary trial period of 6 months, during which the employee can be dismissed on the grounds of the employee's lack of suitability for the work or lack of proficiency or reliability following a 14 day notice. Fixed-term hiring has stricter regulations, and an employee can only be temporarily hired if the work is also temporary, or if the employee is a temporary replacement hire, a trainee, or a participant in an active labor market program.

Union membership in Norway is relatively high compared to other countries in the OECD and the U.S., but has fallen from 58 to 53 percent from 1992 to 2013.[6] Still, virtually all private sector jobs are covered by collective bargaining agreements, and wages and working hours are typically set in accordance with collective agreements between unions and employer associations. Tariff wages at the industry-level are first set centrally, after which wages are supplemented by local adjustments, or wage drift, which is bargained over at the firm level. The two-tier framework is considered a key reason for the highly compressed wage structure in Norway, with comparably low inter-industry wage differentials (see Barth *et al.*, 2014).

3 A Worker Flow Approach to Internal Labor Markets

In this section, we describe our method. Our method identifies internal labor markets and job ladders within firms based on observed worker flows across occupations. In contrast to the traditional view of firms where individuals contract on spot markets, internal labor markets are defined as coherent clusters of jobs where the allocation and pricing of labor are governed by administrative rules (Doeringer & Piore, 1971). These rules define career paths along which workers move from job to job during their careers. We use the observed job-to-job transitions within firms to recover the underlying structure of internal labor markets. We proceed in two steps. First, we identify which jobs form an internal labor market in each firm. Second, for each internal labor market, we identify the hierarchical structure of jobs.

[6] OECD Statistics Trade Union Statistics, Accessed June 14th, 2020

3.1 Internal Labor Markets

To examine and estimate the structure of internal labor markets in firms, we study worker transitions between occupations within firms. The main idea is that observed transitions of workers between positions in a firm reveal the boundaries of internal labor markets. Based on the concept of connected components from graph theory, an internal labor market comprises all occupations that are connected by some path that is established through realized worker flows. Hence, our algorithm identifies sets of interrelated occupations that together form an internal labor market. Importantly, it can separate one internal labor market from another or from single occupations that are filled exclusively with external hires.

3.1.1 Network Components

Our method is closely related to the approach used to examine the internal structure of a particular firm in the seminal contribution of Baker *et al.* (1994). In contrast to that paper, our method is entirely data-driven. Moreover, we can examine the full population of firms and can extend the analysis from management occupations to the entire workforce of each firm. An intuitive example would be a firm in the manufacturing sector that has a large internal labor market for their core production workers, but also employs workers in separate internal labor markets for administrative purposes, or even in single occupations like security personnel or receptionists that are unconnected from the internal labor market. Our method deviates from market-wide approximations of internal labor markets that are defined by observable categories such as establishments (Cestone *et al.* (2019); Giroud & Mueller (2019)) or broad occupational categories (Lazear & Oyer (2004)).[7] We observe a large fraction of job transitions across broad occupational categories and establishments. Hence, simply aggregating jobs into broader (e.g., 2-digit) occupational categories does not capture relevant internal labor markets.[8] A potential limitation of our study is that the detailed occupational data is available only for 8 years. To the extent that transitions between different jobs in the internal market are less frequent than 8 years, we cannot capture connections between occupations with very low turnover.

Specifically, we define internal labor markets within each firm by constructing connected components of occupations. Two occupations are connected if within our sample period a worker transitions between those two occupations. A connected component is a set of occupations where all occupations in the set are connected by some path of transitions but there are no connections to occupations outside the set. We interpret these components as internal labor markets where hiring in one component is independent of hiring in the others. This data-driven approach allows us to identify internal labor markets within the firm. Using the observed matrix of within-firm transitions, connected components can be very easily identified using a simple breadth-first search algorithm that is implemented in standard statistical packages.

[7] Our approach nests these other structures. If the mobility of workers is in fact limited to an establishment or a broad occupational category, that will be the endogenously determined in our model.

[8] In the Norwegian data, 54 percent of all within-firm occupational switches are also transitions across much broader 2-digit occupations. This is in line with recent evidence from the US (Schubert *et al.*, 2020) where 86 percent of all 6-digit occupation switches (across firms) are also 2-digit switches. Moreover, 13 percent of all within-firm job switches in our data are also switches of the establishment.

3.1.2 Handling Measurement Error

A potential concern is that the data might contain measurement error in occupational coding. Although firms are obliged to report their employees' occupation, they have no strategic incentives to do so correctly. Since we are mainly interested in transitions between occupations *within* firms, the issue appears less problematic as we can reasonably assume that the misclassification of occupations is constant within firms. Occasionally, however, due to turnover in HR staff or due to errors in data entry, we may observe implausible transitions between seemingly unrelated occupational codes. To address this concern, we employ a data-driven cleaning procedure that aims to separate plausible from implausible links. Our pruning algorithm iteratively removes workers from the network and assesses whether the removal breaks the network apart. Removing a worker breaks the network apart if the link established by this worker accounts for less than 10 percent of all transitions into or out of the jobs between which the worker switches and, hence, is an exceptional move.

The pruning algorithm is based on the concept of bi-connected components in graph theory and akin to a procedure previously used in a related context by Kline *et al.* (2019) (see Appendix B for a detailed description of the algorithm). In contrast to the leave-one-out components employed in (Kline *et al.*, 2019), we extend the algorithm to allow for rare but true transitions. Intuitively, if the observed network contains one transition from a senior manager position to the CEO position, this transition should not be capped to split the network. If, however, the network contains many accountants and many financial analysts who become financial managers and only one secretary who switches to financial manager, the algorithm breaks the link between the secretary and financial manager position. We call this the leave-X-percent-out procedure and choose 10 percent as our baseline. This proportional method has the additional advantage that it is scale-invariant, i.e., it yields the same classification of internal labor markets in two firms with the same structure but a different number of workers and transitions.[9]

3.2 Job Ladders

In the second step, we identify job ladders within internal labor markets by tracing the hierarchical structure of job flows. We rank all occupations within an internal labor market based on the relative flows between them. The basic intuition is simple: More observed transitions from occupation b to occupation a than vice-versa imply that occupation a is ranked higher than b. Using this simple logic, we overcome the fundamental challenge that typically linked employer-employee data does not directly allow us to observe job hierarchies.[10] While the literature in personnel economics has embraced the idea of using worker flows to elicit job hierarchies (e.g., Baker *et al.*, 1994), the tedious process of collecting data and evaluating the flows between job titles has limited the existing evidence to case studies from particular firms. To extend the method of BGH beyond one firm, we introduce a minimum violations ranking approach. This approach

[9] We thank our discussant Thomas Lemieux to point out this idea.

[10] Several data sets and standard occupational coding include very rough classifications of occupations into levels. These rough levels can typically explain a relatively large portion of wage variation (see e.g., Caliendo *et al.* (2015a); Bayer & Kuhn (2019); Lazear & Oyer (2004)). Our algorithm, however, has several advantages over such predefined categories. First, it is firm-specific and therefore able to capture even subtle forms of firm heterogeneity. Second, it is much more detailed and allows us to examine promotions, demotions, and the underlying incentives on a much finer scale. Finally, if all job transitions follow such predefined categories our algorithm will trace out exactly those categories as the relevant hierarchical structure.

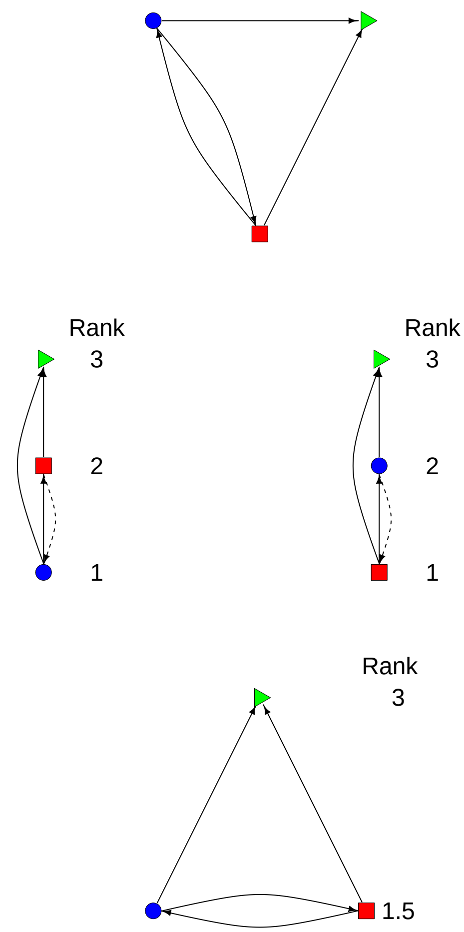
orders occupations in an internal labor market such that there are as few transitions downwards (i.e., reverse) the hierarchy as possible. In this section, we describe the ranking algorithm and describe how we group ranks into levels of hierarchy.

3.2.1 Minimum Violation Ranking

The ranking algorithm orders occupations within each ILM based on relative flows between them. While the idea that *a* ranks higher than *b* if more workers move from *b* to *a* than in the opposite direction is simple and intuitive, the actual estimation is more complicated because the observed relations between occupations are not transitive. We therefore use an algorithm that ranks occupations within each internal labor market such that the number of transitions towards lower steps on the occupational job ladder is minimized (**minimum violation ranking**). The resulting ranking represents ILM-specific hierarchies among jobs. The ranking algorithm is based on the fraction of links that are upward moves along the job ladder, i.e. the ranking of the target occupation is higher than that of the source. We maximize this fraction over all possible rankings following the procedure in Clauset *et al.* (2015). Starting from an initial ranking where occupations are ranked according to the number of outbound transitions, we converge to the optimal ranking by repeatedly swapping ranks of a randomly chosen pair of occupations and accepting swaps where the associated new ranking has the same or a higher number of upward moves (or, equivalently, a lower number of ranking "violations"). Since there are potentially several equally plausible rankings with the maximum possible fraction of upward moves, we sample optimal rankings from the set of permutations with the maximal fraction of upward moves. Our results are then averaged rankings over the sampled sets while uncertainty around the estimated ranking can be measured by the distribution of ranks across optimal rankings.

Figure 1 illustrates the logic of our algorithm. Suppose the ILM consists of three occupations—green, blue, and red—that are connected as shown in the top panel of the figure. Two possible rankings share the same minimum number of one violation (see the two middle panels). In both cases, the green triangle is the highest-ranked occupation. The red square and the blue circle, however, are each ranked second in one case and ranked lowest in the other case. In order to find a consensus ranking, the MCMC algorithm will therefore—after having converged to the minimum number of violations—collect many (random) samples from the set of possible rankings with the lowest number of ranking violations and subsequently average the ranks from all samples. The lower panel of Figure 1 shows that in the consensus ranking, the blue circle and the red square both receive rank 1.5 as we expect them both to be ranked 1 and 2 in half of the samples. Running the algorithm several times also provides us with a measure of uncertainty around the estimated rankings. Due to the stochastic nature of the algorithm, the consensus ranking will vary across runs. We compute the standard deviation of ranks across these runs as our measure of uncertainty.

Figure 1: Two potential minimum violation rankings



Notes: These figures illustrate the logic of our ranking algorithm. Suppose the component consists of three occupations—green, blue, and red—that are connected as shown in the top panel of the figure. The middle two figures illustrate two possible rankings that share the same minimum number of one violation. In both cases, the green triangle is the highest-ranked occupation. The red square and the blue circle, however, are each ranked second in one case and ranked lowest in the other case. The lower panel shows that in the consensus ranking the blue circle and the red square both receive rank 1.5 as we expect them both to be ranked 1 and 2 in half of the samples.

3.2.2 Job Levels

After estimating the minimum violation ranking, we group occupations into hierarchy levels. This allows us to summarize the hierarchical structure of ILMs by a small number of relatively homogeneous sets of occupations that are comparable across firms. To do so, we cluster occupations with similar ranks into the same level of hierarchy using a k-means clustering algorithm (see, e.g., Bonhomme *et al.* (2019)). In contrast to the clustering procedure in Bonhomme *et al.* (2019), we cluster occupations based on a single dimension: the estimated rank. This enables us to employ a stable dynamic programming algorithm that guarantees optimality and reproducibility (Wang & Song, 2011). Given a number of clusters K, the k-means algorithm assigns each occupation to a hierarchy level such that the sum of the within-level squared distances in estimated ranks is minimized. The choice of the number of hierarchy levels for each internal labor market is determined by the uncertainty in the rank estimation. In particular, following the suggestion in Bonhomme *et al.* (2019) we choose K such that the value of the k-means objective function is at least as low as the average standard deviation of the estimated ranks. Details are provided in Appendix B.

3.3 Illustrative Example

We illustrate our method based on an example firm to which we apply the algorithms described above. The firm is in the manufacturing industry, has four plants, and consists of a large ILM with 22 occupations and 2,412 worker-year observations. After applying our pruning procedure, the firm has five single occupations that are not connected to the main ILM. Column (5) of Table 2 shows these numbers and emphasizes that 96 percent of all worker-year observations are part of the ILM.

Internal Labor Market. Panel A of Figure 2 depicts the structure of the ILMs in the sample firm. The single occupations are Lawyer, Financial Consultant, Key Account Manager, Webmaster, and Business Analyst, and are not interacting with the main internal labor market of the firm.[11] Instead, these occupations are hired externally, and when workers leave these positions they move to somewhere outside the sample firm. These are occupations with skills presumably tied to the occupation rather than to the core business of the firm. In terms of the individual wage component, measured by the average AKM person fixed effect in the occupation, there are single occupations both on a higher and lower level than the average in the ILM. The size of the circles in the figure is proportional to the number of workers and emphasizes that the single occupations are small.

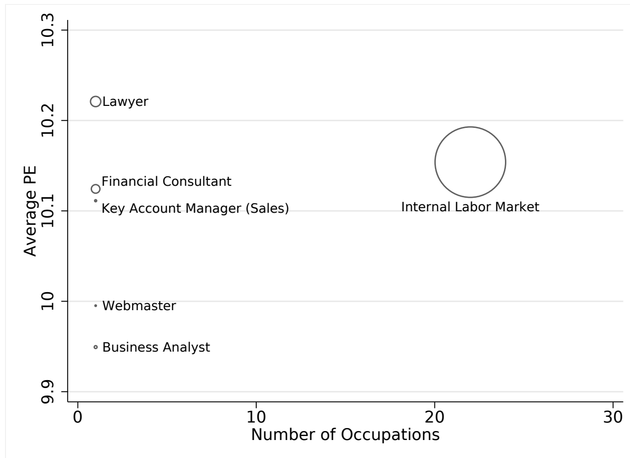
Ranks and Hierarchy Levels. Panel B of Figure 2 zooms into the ILM of the sample firm and shows how the estimated ranks are grouped into ten hierarchy levels using our k-means clustering algorithm. The vertical bars indicate 95 percent confidence intervals for the estimated rank of a given occupation. We note that the clustering is based on the mean estimated rank only and the number of clusters is determined by the overall uncertainty in estimated ranks. Nevertheless, most levels of hierarchy are clearly statistically distinct from each other while positions within the same level have mostly overlapping confidence intervals. In the middle of the hierarchy, there is more noise around the classification. The classification of occupations into hierarchy levels appears to reasonably reflect the hierarchical structure of the job titles. We find several director and senior advisor positions towards the top of the hierarchy while the lower end of the hierarchy is populated by consultant, advisor, and research positions.

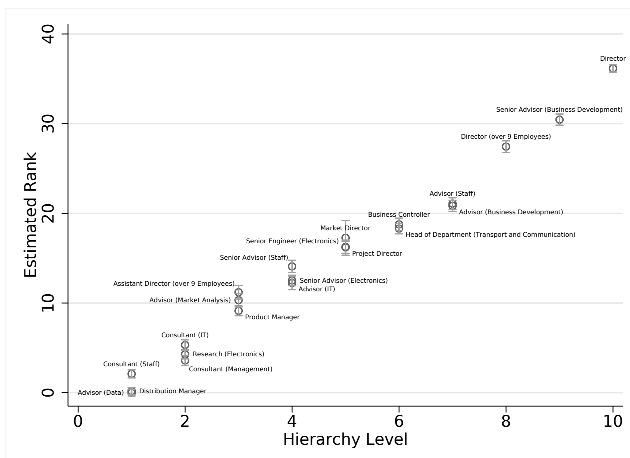
Job Ladders. Panel C of Figure 2 shows how the job ladder structure of our sample firm relates to the person wage fixed effect from our AKM decomposition. The figure shows a positive relation between the hierarchy level of each occupation and the average AKM person effect in that occupation. As before, there is somewhat more variation in the middle of the hierarchy structure.

[11] Either there does not exist any observed link to the larger internal labor market, or their connections are rare exceptions that are removed by our cleaning procedure.

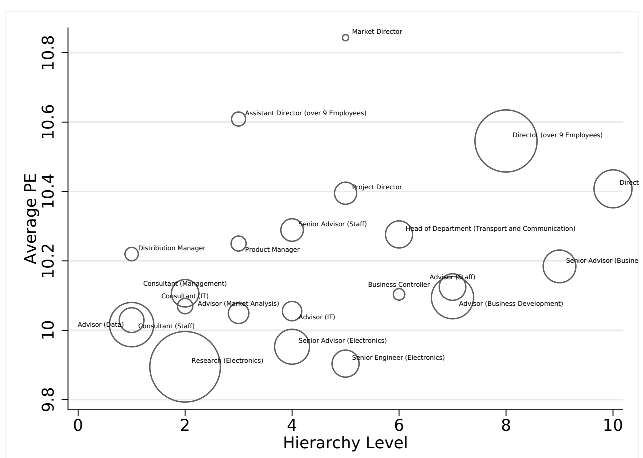
Figure 2: Internal Labor Market and Job Ladders in a sample firm

(a) ILM and single occupations





(b) Ranks and Hierarchy Levels



(c) Hierarchy Levels and AKM Person Effects

Notes: Panel A illustrates the internal labor market structure of an example firm. Panel B shows the estimated ranks of the occupations in the ILM and how they are assigned to levels of hierarchy. Panel C shows the relation of the hierarchy levels and AKM person effects. The larger the circle, the more individual workers are employed in the internal labor market.

4 Main Findings

This section applies the methods introduced in Section 3 to our sample of Norwegian firms and presents the main evidence on the structure of internal labor markets and job ladders. We validate that the estimated structures reflect the main properties of internal labor markets established in the theoretical literature and compare our findings to the existing empirical evidence.

4.1 Internal Labor Markets

Table 2 reports summary statistics on our full set of internal labor markets. In total, there are 2089 firms that comprise about 5,500 components. While the average firm has 2.6 distinct network components, Table 2 shows that a large fraction of the firms has one large internal labor market where workers move between occupational titles. Additionally, about a quarter of the firms in our sample have a second, smaller ILM and the average firm uses 1.2 job titles that are not connected to the rest of the firm (after our pruning procedure).[12] The vast majority of workers (97 percent) and occupations (81 percent) are part of the largest internal labor market in their firms. The average internal labor market consists of 6.9 occupations. On average, the largest ILM covers 891 worker-year observations, while there are only 44 worker-year observations in smaller ILMs and 16 worker-years in the average single occupation. However, there is quite a substantial variation across firms.

[12] As we are interested in internal mobility, we refer to network components with two or more occupations as internal labor markets (ILM) and to components with only one occupation as single occupations.

Table 2: Internal Labor Markets and Single Occupations

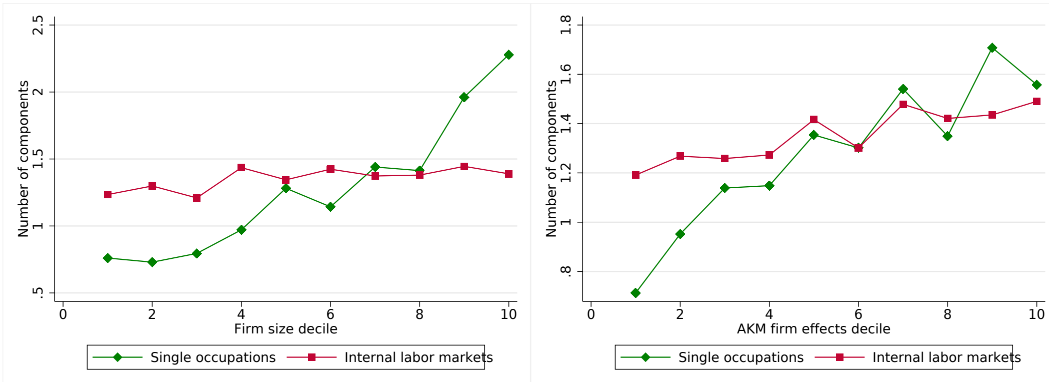
	Leave-out Components				Sample Firm
	Avg.	25th pct	Median	75th pct	Avg.
Number of Components	2.630	1	2	3	6
Internal Labor Markets	1.35	1	1	2	1
Single Occupations	1.28	0	1	2	5
Share worker-years largest ILM	0.97	0.90	0.99	1	0.96
Share occupations largest ILM	0.81	0.67	0.83	1	0.81
N x T largest ILM	890.8	162	355	878	2412
N x T small ILMs	43.6	10	20	45	.
N x T single occupations	15.5	2	5	12	18.8
# Occupations ILM	6.88	2	4	8	22
# Firms	2089				1

Notes: This table reports characteristics of the internal labor markets. The sample is restricted to private sector firms with at least 30 employees and at least 10 job transitions within the firm during the period 2006-2014. We further restrict our sample to firms with at least 15 external hires over the period. The sample of workers includes every full-time employed male and female aged 20 to 61. If a person has multiple employers over the year, the data includes the employer in the last month observed; and, if the person has multiple employers in the last month, we select the employer with the highest total earnings in a particular year. We also exclude workers who have a period of more than 6 months in non-employment between an external and internal move. The sample is further restricted to firms where we can identify the organizational structure. This is limited to firms with less than 200 occupational links. Components are connected sets of occupations using our leave-out pruning procedure. Internal labor markets are components with at least two occupations. N x T refers to the number of worker-year observations, in the largest internal labor market, secondary labor markets, and single occupations, respectively.

Figure 3 relates the number of internal labor markets and single occupations per firm to firm size and to the estimated AKM firm fixed effect. The graphs paint a consistent picture. The number of internal labor markets per firm is a little above one and roughly constant across firm size (measured by the number of worker-years) and across firms with different pay premia (represented by the firm fixed effect from the AKM decomposition). The number of single occupations is increasing with firm size and firm fixed effects. This suggests that larger and higher-paying firms also rely on additional functions that are hired predominantly on the external labor market and are not part of the internal labor market.[13] The vast majority of workers and occupations, however, are part of the core internal labor market even in large firms. Based on this evidence, we concentrate on the largest internal labor market of each firm in the remainder of this paper.

[13] Figure A1 in the Appendix shows that firms with a broader variety of occupations, measured by the Herfindahl index of employment shares, are also characterized by a higher number of single occupations.

Figure 3: Components by Firm Size and Firm Fixed Effect



(a) Panel A: Firm Size (b) Internal Labor Markets and AKM Firm Effect

Notes: Panel A illustrates the number of ILMs per firm, and the number of single components (with only one occupation) per firm by firm size. Panel B illustrates the number of ILMs per firm, and the number of single components (with only one occupation) per firm by AKM firm fixed effects. The sample consist of private sector firms over the period 2007-2014. The sample is described in Table 2.

4.2 Job Ladders

We now turn to the hierarchical structure of internal labor markets. We focus on the largest internal labor market of the firm and restrict ourselves to internal labor markets with at least 2 occupations. This corresponds to a sample of 2,076 internal labor markets.[14] Overall, we document substantial variation in the hierarchical structure of internal labor markets. About 40% of our sample of firms has three or four levels, while the longest job ladders have up to 19 hierarchy levels. In the following, we classify ILM's by the number of hierarchy levels (i.e., the number of rungs of the internal job ladder).

Firm characteristics by firm type

Table 3 shows firm characteristics by the number of hierarchy levels. As firm size increases and the number of occupations expands, hierarchies get longer. Partly, this is mechanical, because the number of hierarchy levels can only increase if there are more occupations. But we also see that the average number of occupations by level is increasing in the hierarchy, from 1.4 occupations per hierarchy level in firms with two levels to 2.8 occupations per level in firms with 8 or more levels.

Perhaps not surprisingly, the probability of promotions increases with a growing number of hierarchy levels, along with the probability of demotions. The probability of being promoted in a single year is around 6 percent for workers in firms with longer hierarchies. This means it takes on average almost 17 years to be promoted to a higher level. Relative to promotions, demotions are observed less frequently; about one percent of the average workers experience a demotion in a given year.[15]

Next, we examine the characteristics of the workforce in ILMs with a different number of hierarchy levels. The share of workers with management positions is increasing in the length of the job ladder, which

[14] In this section, we refer interchangeably to the ILM as internal labor market or firm.

[15] The low frequency of promotions and the existence of demotions, suggest that the job ladders we identify do not correspond to automatic movements along a pay scale combined with an occupational upgrade, where all workers would typically move more regularly.

indicates that firms with longer hierarchies also have more complex supervisory structures. Worker tenure is increasing in the number of hierarchy levels as well, which confirms that workers tend to stay longer in firms with longer job ladders.

Average wages increase by about two log points from flat firms with only one hierarchy level to firms with the longest hierarchies. This average wage change comes with a one log point increase in both the average firm effect and the average worker effect, which indicates some sorting. Firms with longer hierarchies are paying higher wages to all their workers, but they also employ more highly skilled workers.

Table 3: Descriptive Statistics

	Number of Levels of Firm							
	1	2	3	4	5	6	7	8 +
Size	33.5	52.0	83.2	143.4	222.2	287.3	377.3	594.8
Growth in Size	0.76	0.72	0.68	0.90	0.72	0.58	0.30	0.52
Number of Occupations	2.30	2.83	4.62	7.19	10.4	13.3	16.5	28.1
Number of ILMs and single occupations	1.89	2.11	2.25	2.50	2.77	3.11	3.36	5.11
Number of Single Occupations	0.80	0.80	0.92	1.16	1.35	1.71	1.96	3.49
Number of Years Obs.	7.35	7.18	7.51	7.53	7.58	7.55	7.77	7.89
Prob of Promotion	0	0.035	0.037	0.044	0.047	0.056	0.059	0.062
Prob of Demotion	0	0.015	0.0095	0.010	0.011	0.012	0.015	0.014
Share Female	0.22	0.31	0.33	0.33	0.30	0.31	0.29	0.29
Share Management	0.065	0.071	0.059	0.057	0.063	0.10	0.087	0.095
Wage	10.4	10.5	10.5	10.6	10.6	10.7	10.7	10.7
Wage Growth	0.022	0.022	0.023	0.023	0.023	0.024	0.027	0.025
Person FE	9.76	9.79	9.79	9.80	9.78	9.87	9.82	9.84
Firm FE	0.49	0.54	0.54	0.57	0.58	0.61	0.62	0.64
Tenure	62.0	66.3	64.6	66.7	70.9	72.1	77.9	78.1
Number of Firms	54	636	496	325	190	114	77	184

Notes: The table displays characteristics by firm types. The sample consist of 2076 firms where the largest component has more than one occupation. Firm types are defined by the number of levels. 8 + levels refer to firms with 8 – 19 levels. Component level variables refer to the largest component in the firm. Time varying variables are measured in the last year when the firm is observed or in the year where the variable is highest. The sample is restricted to private sector firms with at least 30 employees at some point during the period 2007-2014, and that we observe at least 10 job transitions within a firm (or at least two distinct occupations). We further restrict our sample to firms with at least 15 external hires over the period. The sample of workers includes every full-time employed male and female aged 20 to 61. If a person has multiple employers over the year, the data includes the employer in the last month observed; and, if the person has multiple employers in the last month, we select the employer with the highest total earnings in a particular year. We also exclude workers who have a period of more than 6 months in non-employment between an external and internal move. The sample is further restricted to firms where we can identify the organizational structure. This is limited to firms with less than 200 occupational links. Wage is at nominal levels. Observations are worker-year.

Characteristics of hierarchy levels by firm type

Some of the firm characteristics indicate that the estimated hierarchies capture the main concepts of internal labor markets with job ladders. In order to get a more detailed picture, we now turn to descriptive statistics of separate hierarchy levels in Figure 4.[16] The graphs in this figure show averages by hierarchy level for different firm types. We first investigate employment, in terms of worker-year observations, by hierarchy

[16] We group firms with longer ladders into one category with 8 or more hierarchy levels, for expositional clarity. But the patterns hold for longer job ladders too.

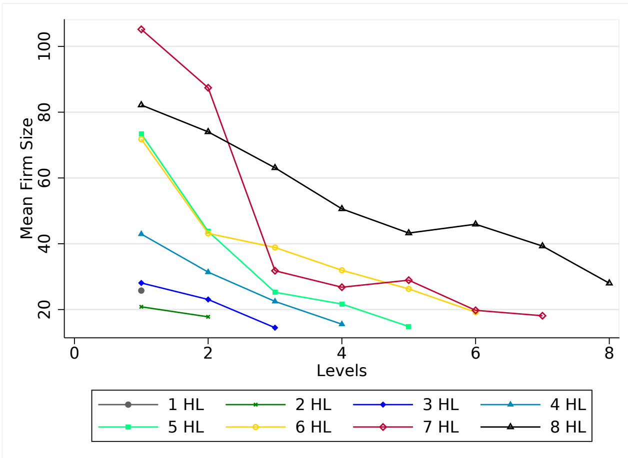
level in Figure 5a. We have already seen that firms with longer hierarchies are larger, which is reflected in the graph. But it also shows that the number of workers is decreasing in the hierarchy level across firm types. This means that hierarchies have a pyramid structure, a finding also pointed out by Baker *et al.* (1994). We consistently find the pyramid structure of the average firm across all firm types.

Figure 5b shows that the average number of occupations across levels is hump-shaped, again consistently across firm types. At the entry-level there are fewer occupations and specialization tends to increase towards the middle of the hierarchy, before it declines at the top hierarchy levels. This pattern is also consistent with the organizational chart of the firm in Baker *et al.* (1994).

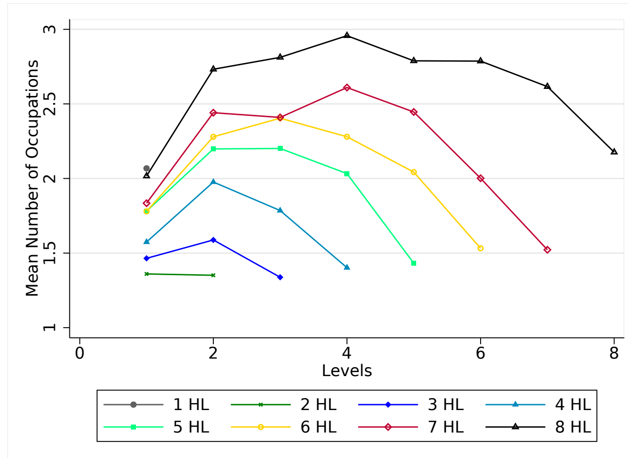
According to Doeringer & Piore (1971), workers enter an internal labor market at entry positions and then climb up the job ladder. This strict form of internal labor market is typically found in bureaucratic organizations (see, e.g., Bertrand *et al.*, 2020). For private sector firms, the empirical literature on job ladders generally documents less strict ladders with a declining share of external hires over hierarchy levels (Baker *et al.*, 1994; Lazear & Oyer, 2004). Figure 5c confirms this finding for the Norwegian market. The share of external hires is about 90 percent in all firms at the lowest hierarchy level and it declines to around 55 percent at top hierarchy levels. Interestingly, the figure shows an almost identical development in the share of external hires across firm types, which suggest that levels are comparable across firms, and even across firms with a different number of hierarchy levels.

Finally, Figure 5d shows that the share of workers in management occupations increases with the hierarchy level. This occupational category includes mid-level and top-level management, for example, managers in a retail shop as well as the CEO and the executive board. The management share is close to zero at the entry-levels in all firm types and increases over the hierarchy structure. This pattern is in line with the intuition that task complexity and responsibilities increase as workers move up the job ladder. Interestingly, both the low shares of management at the lowest levels and the increasing pattern of management shares with level is similar across firm types.

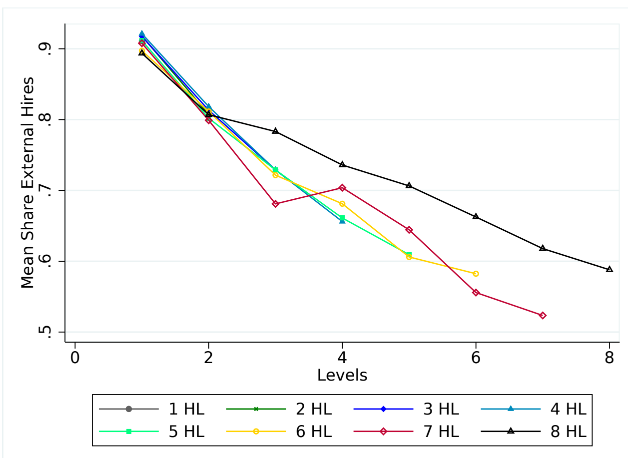
Figure 4: Empirical Job Ladders



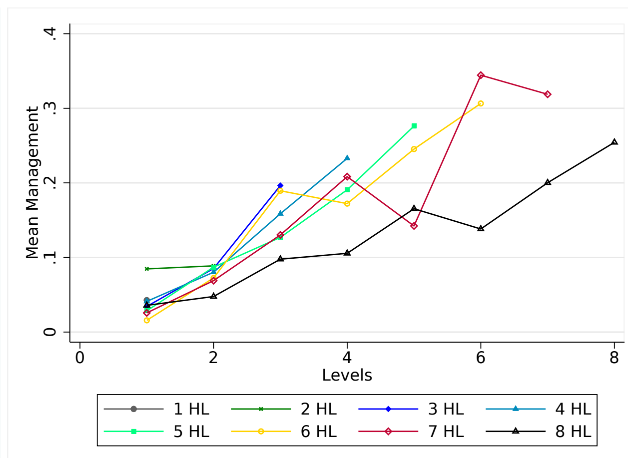
(a) Employment Level



(b) Number of Occupations



(c) Share External Hires



(d) Share Management

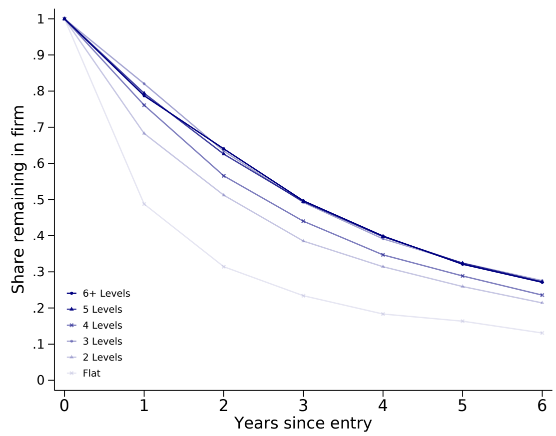
Notes: This figure illustrates the key characteristics of our empirical job ladders by firm types. HL in the figure refers to the number of hierarchy levels. We group ILMs with 8 or more hierarchy levels into one category and plot the averages for all ILMs in this group. We consider the largest ILM per firm of our sample of private sector firms over the period 2007-2014. The sample is described in Table 3.

Worker careers.

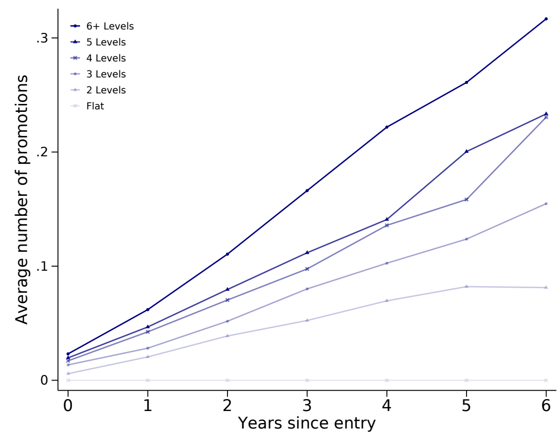
The patterns shown in Figure 4 provide a strong indication that the estimated hierarchy structures capture common properties of internal labor markets. According to Doeringer & Piore (1971) a key concept to internal labor markets is that workers have careers within the firm. We proceed by examining individual workers' career paths in firms with a different number of hierarchy levels. To track the career within a firm, we construct a sample of workers who enter one of our firms in 2007 or 2008 at any level of the hierarchy. We then follow the promotion and demotion dynamics of these workers in the same firm over the following six years.

The Panels of Figure 5 document the average career paths of the new entrants. Figure 6a examines how the duration of the employment spell varies with the firm type. We observe quite a lot of turnover, especially in the first 2 to 3 years, and on average only 30 percent of entrants remain in the same firm for 6 years. The

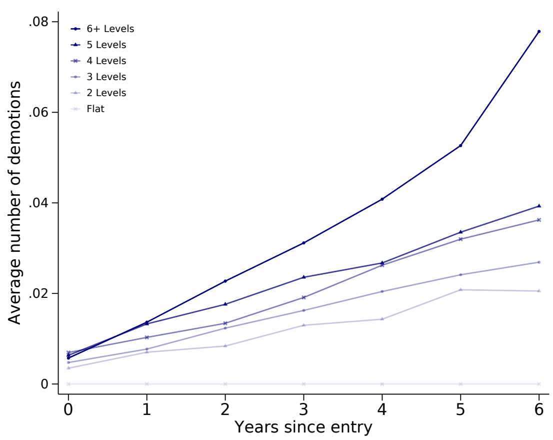
Figure 5: Careers



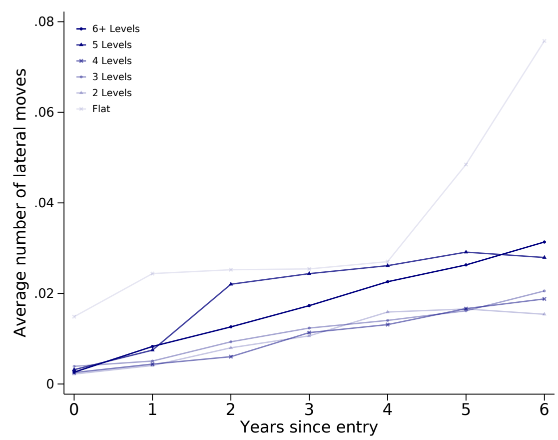
(a) Share of workers remaining in firm



(b) Number of internal promotions



(c) Number of internal demotions



(d) Number of lateral job changes

Notes: Figure 6a illustrates the attrition of workers since entering the firm. Figure 6b illustrates the total number of moves in the same firm. A move is when a worker switches job title. Figure 6c illustrates the total number of promotions in the same firm. A promotion is when a worker moves to a higher level in the firm. The sample is restricted to workers who join the firm from 2007 to 2009 (see further details of the sample restrictions in Table 3).

figure also shows that exit probabilities vary by firm type with a clear ranking. Firms with longer job ladders potentially offer better prospects of promotions and are able to keep workers longer. By contrast, firms with flat hierarchies experience the highest level of churn.

The next panels decompose the career dynamics of workers who are moving between positions in the same firm into three categories. Panel 6b shows a very clear ranking in the cumulative number of promotions across firm types. While there are no promotions, by definition, in flat firms, up to 30 percent of workers who remain in a firm with at least six levels for 6 years experiencea promotion. The graph also indicates the probability of being promoted slightly increases over the years. But we cannot distinguish whether this is due to dynamic selection or an increase in the probability of promotion with tenure. Panel 6c, shows the corresponding graph for demotions. Across all firm types demotions are much rarer than promotions and happen about one fifth as often to workers still in the firm after 6 years.[17] Finally, Panel 6d shows the cumulative number of lateral moves (i.e., moves within the same hierarchy level). Similar to demotions, lateral moves are relatively rare, and tend to happen slightly more often in larger firms with longer job ladders.

4.3 Internal Labor Markets and Wages.

In internal labor markets, wage setting should be tightly linked to the job hierarchy. On the one hand, the internal labor market makes the firm somewhat independent from competitive market forces, and on the other hand, firms can use the internal hierarchy structures to design promotion systems as incentive or allocation mechanisms. One popular model of promotions is the tournament model Lazear & Rosen (1981), where workers face increasing wage incentives to reach higher hierarchy levels. This model predicts discontinuous wage increases at each hierarchy level and increasing wage spreads between hierarchy levels the higher the worker climbs in the hierarchy.

We confirm that the estimated job ladders in Norwegian firms reflect the predictions from this theory. Figure 7a shows that average log wages are increasing by hierarchy levels. The upward pattern appears somewhat steeper for firms with shorter hierarchies. But importantly, the profile for each firm type is almost linear in mean log wages. In approximation, this implies that the wage spreads between levels are increasing over the hierarchy, or the prices for winning the tournament are increasing towards the top of the hierarchy. This result extends Baker *et al.* (1994) and other papers that provide empirical support for an increasing spread of average wage levels along job hierarchies.

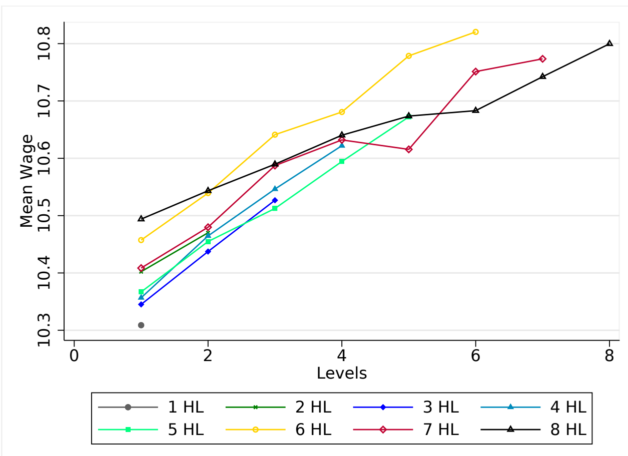
An implication of wage setting models with promotions is that the wage premium at higher levels motivates workers to remain in the firm. The incentive structure should thus lead to longer tenure in firms with job ladders, especially among workers at higher parts of the job ladder, and to sorting of higher ability types to the upper part of the hierarchy, because they can climb upwards faster. Alternatively, theories of job assignment also imply that firms use the hierarchy to assign workers with higher ability towards higher levels (see, e.g., Gibbons & Waldman, 1999a). We can assess whether job ladders reflect sorting of high-ability types to higher hierarchy levels by investigating firm tenure and the worker fixed effect from AKM wage

[17] This analysis offers a different statistic of the performance of our ranking algorithm. The relationship of promotions to demotions confirms the performance of our ranking method, which had the aim of minimizing downward career moves.

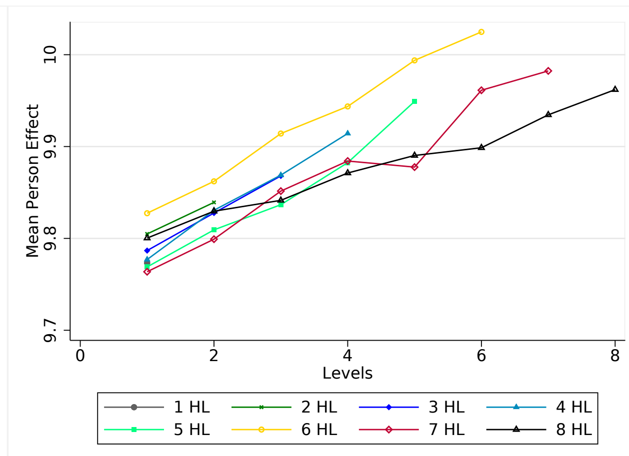
decompositions as a proxy for ability.

Figure 7b shows the average person fixed effects from the AKM wage decomposition by hierarchy level. The fixed effects can be seen as measures for worker characteristics that are transferable across firms. Similar to wages, person fixed effects follow parallel and upward sloping patterns in hierarchy levels across all firm types. In entry-level jobs, the average person effect is substantially lower than at higher hierarchy levels. This suggests that higher worker types move up the job ladder more quickly and this mechanism generates a sorting pattern. In Figure 7c, we plot average worker tenure by hierarchy level for different firm types. It shows that although tenure is generally increasing in the hierarchy at all levels, there is are also increasing trends over hierarchy levels within firm types.

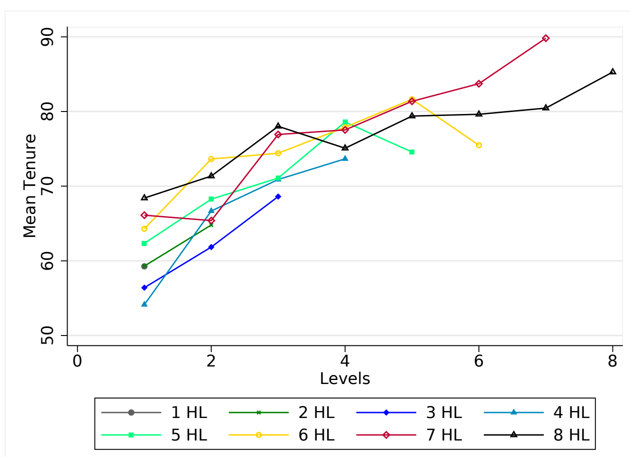
Figure 6: Wage and Promotions



(a) Log Wages



(b) AKM Person fixed effects



(c) Tenure

Notes: This figure illustrates the key characteristics for our empirical job ladders by firm types. HL in the figure refers to the number of hierarchy levels. We group ILMs with 8 or more hierarchy levels into one category and plot the averages for all ILMs in this group. We consider the largest ILM per firm for ou sample of private sector firms over the period 2007-2014. The sample is described in Table 3.

Individual Wage Regressions.

Baker *et al.* (1994) show that hierarchy levels explain about 70 percent of the cross-sectional variance of wages within the firm (and substantially more than typical Mincer regressions). This finding is echoed in a recent paper by Bayer & Kuhn (2019) who use five rough levels of job hierarchies reported in German data. We examine whether this holds on average and explore the distribution across different firms.

Table 4 reports regression results for a pooled regression of log wages on human capital measures and hierarchy level indicators. The first column reports the explained variance of the benchmark model with fully interacted firm and year fixed effects with an adjusted R2 of 0.48. The second and third columns show that controlling for a quadratic polynomial in tenure and including dummies for years of schooling add about 2 percentage points to the explained variance. The fourth column shows the full specification with 19 hierarchy level dummies. Comparing the adjusted R2s reveals that the hierarchy structure explains about as much of the variance in log wages as the tenure and human capital variables together. With an adjusted R2 in column 4 of 0.52, there still appears to be a substantial individual component to wage setting.

The coefficients on the hierarchy level dummies, also plotted in Figure 7, confirm that promotions give an extra push to wage growth beyond the wage increases due to seniority. Within the same firm and holding individual characteristics fixed, the average wage in the top hierarchy level is about 40 percent higher than wage at entry level. The increase in the (log) promotion bonus is almost linear at least from level 4 onwards. This again confirms the hypothesis of increasing spreads between hierarchy level from tournament theory. It also implies that the gap in wage growth between workers in firms with longer hierarchies and to workers in firms with shorter hierarchies, leads to increasing income inequality over the life cycle.

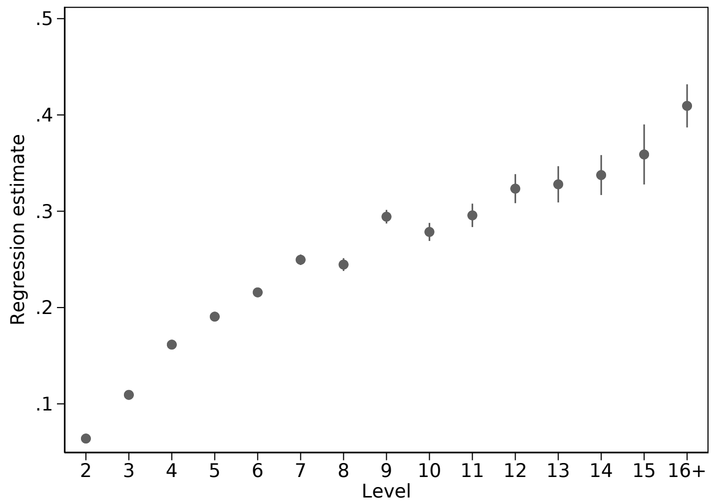
Table 4: Effects of Human Capital and Hierarchical Level on Current Salary - all firms.

	(1)	(2)	(3)	(4)
	ln wage	ln wage	ln wage	ln wage
Tenure		0.00196***	0.00209***	0.00203***
		(0.0000167)	(0.0000166)	(0.0000159)
Tenure2		-0.00000457***	-0.00000478***	-0.00000470***
		(4.97e-08)	(4.93e-08)	(4.71e-08)
Education:				
10-13 years education			-0.00729***	-0.00737***
			(0.00207)	(0.00198)
14-16 years education			0.0968***	0.0758***
			(0.00222)	(0.00212)
Level:				
Level 2				0.0640***
				(0.00128)
Level 3				0.109***
				(0.00152)
Level 4				0.161***
				(0.00185)
Level 5				0.191***
				(0.00222)
Level 6				0.216***
				(0.00239)
Level 7				0.250***
				(0.00277)
Level 8				0.245***
				(0.00338)
...				⋮
Level 16-19				0.409***
				(0.0114)
Observations	1,840,592	1,840,592	1,840,592	1,840,592
R2	0.482	0.494	0.501	0.519
Adjusted R2	0.478	0.490	0.497	0.515
Sex and age FE	Yes	Yes	Yes	Yes
Firm x year FE	Yes	Yes	Yes	Yes

Standard errors in parentheses are robust and clustered at the individual level * p<0.05, ** p<0.01, *** p<0.001.

Notes: The sample of workers is restricted to full-time employed in the period 2007-2014 in private sector firms (see further details of the sample restrictions in Table 3). Column (1) displays the output from a regression of log wages on age dummies, tenure and tenure squared. Column (2) displays the output from a regression of log wages on age dummies, tenure, tenure squared and education dummies (years of schooling). Reference level of years of schooling is 0-9. Column (3) displays the output from a regression of log wages on age dummies, tenure, tenure squared, education dummies and dummies for each level in the hierarchy. The reference level in the organizational structure is 1.

Figure 7: Mincer Regression - full sample.



Notes: This figure presents the regression coefficients from Table 4, column 4. The sample of workers is restricted to full-time employed in the period 2007-2014 in private sector firms (see further details of the sample restrictions in Table 3). Estimate for level 16 includes levels 16-19. Vetical bars represent 95 percent confidence intervals. Standard errors are clustered at the individual level.

Comparison to Baker, Gibbs and Holmstrom

Our empirical analysis shows that the hierarchical structure obtained by our data-driven method reflects the main characteristics of internal labor markets put forth in the literature. Extending the hand-curated approach of Baker *et al.* (1994) based on worker flows to a fully automated procedure allows us to examine the hierarchical structure of a broad variety of firms in detail.[18]

Our approach identifies an average structure that is similar to the one in BGH, even when we include all types of firms and all workers (and not only management job titles). The empirical results are summarized in Panel A of Table 6. The table shows regression coefficients from a simplified specification where we regress various outcome variables on a quadratic polynomial in the hierarchy level. Additionally, we control for firm and year fixed effect. An increase in the hierarchy level is associated with lower employment, a hump-shaped relation to the number of occupations, a decrease in the share of external hires, and an increase in the share of management, in log wages, and AKM person fixed effects.[19] The detailed graphs in Figures 4 and 6 confirm that the polynomial function is a reasonable approximation of the data.

In the following section, we assess robustness of our empirical findings with regards to several choices in our algorithms.

[18]BGH classified job titles into levels by relying on information about moves between job titles. They initially focused on fourteen titles that each represented at least 0.5 percent of employee-years. They then added the code for Chairman-CEO, and two other titles observed in moves from their original titles to Chairman-CEO to fill in the job ladder to the top of the organization. The lowest level was identified by the hiring patterns: It consisted of job titles exclusively filled externally, who then later moved into job titles at higher levels. The next levels were determined by manually minimizing the rank reversal: Most moves other than stays or exits from the lowest level went to six other titles. These six job titles were only trivially filled by workers coming from other titles where external hires were much less important. This procedure was continued until all job titles were assigned a level.

[19]Panel B of Table 6 shows that very similar results also hold true for our sample firm detailed in Section 3.3.

5 Robustness

In this section, we examine the extent to which our method is robust to making different assumptions about the noise in the data or to using different clustering methods. Furthermore, we assess whether the inferred hierarchy structure is stable over time.

5.1 Sensitivity to Occupational Coding

In our main analysis, we account for potential measurement error in occupational coding by using a pruning algorithm that removes misclassified occupational links. We believe that this is a conceptually important step that helps to distinguish workers that are part of the ILM from those in single occupations that are hired only externally. In this section, however, we show that our findings are robust to omitting the cleaning step and using the unadjusted connected components where we have not removed any links from the networks. Table 5 reports summary statistics for the unadjusted components. The vast majority of firms are still characterized by a single large ILM. In comparison to our baseline (2), there are significantly fewer single occupations in firms and a slightly higher share of workers and occupations is part of the ILM.

Table 5: Internal Labor Markets and Single Occupations

	Unadjusted Components			
	Avg.	25th pct	Median	75th pct
Number of Components	1.56	1	1	2
Internal Labor Markets	1.18	1	1	1
Single Occupations	0.38	0	0	1
Share worker-years largest ILM	0.99	0.98	1	1
Share occupations largest ILM	0.93	0.87	1	1
N x T largest ILM	931.6	185	391	924
N x T small ILMs	39.4	8	19	41
N x T single occupations	14.9	3	6	13
# Occupations ILM	8.65	3	5	10
# Firms	2089			

Notes: This figure documents the characteristics of the internal labor markets when using the unadjusted ILMs. The sample is described in Table 2.

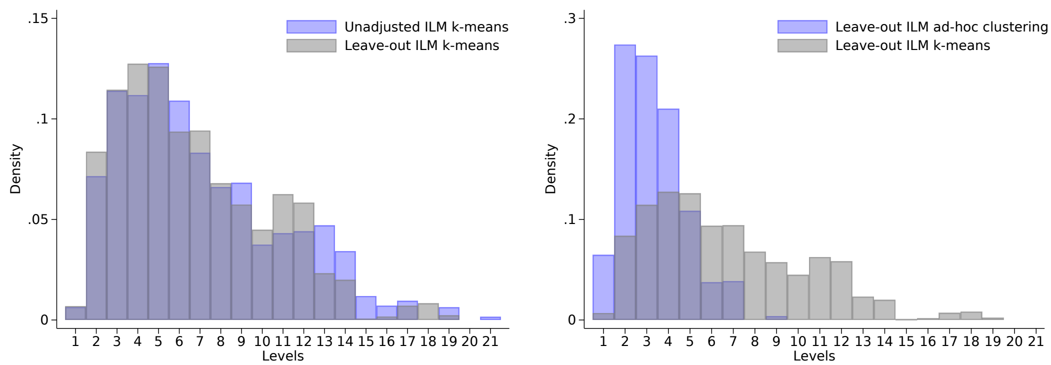
We now turn to the question of whether including these single occupations in the main ILM affects the structure of hierarchy and our empirical results. We re-estimate the hierarchy structure using the unadjusted ILMs that ignore potential misclassification of occupational links. The histogram in Figure 8a shows the worker-year weighted distribution of the number of hierarchy levels in the largest ILM of each firm. Reassuringly, we find that the maximum number of levels is similar, and the proportion of firms with different levels is stable. This suggests that including single occupations into the main ILM does materially affect the structure of the firm.

Panel C of Table 6 shows that also our empirical findings are very robust to omitting the cleaning step. We continue to find a negative relation between employment and hierarchy level, a hump-shaped relation to the number of occupations, less external hires and less managers on higher levels of hierarchy. The relationship between wages and AKM person fixed effects and hierarchy levels remains virtually unchanged.

Figure 8: The Distribution of Job Ladders

(a) Leave-out ILM versus Unadjusted ILM

(b) K-means clustering versus ad-hoc clustering



Notes: This figure shows the worker-year weighted histogram of the number of hierarchy levels per firm. In Figure 8a, grey bars represent our baseline method where we apply the pruning algorithm, and the blue bars represent the unadjusted components. In Figure 8b, grey bars represent our baseline method, and the blue bars represent the ad-hoc clustering method. The sample is described in Table 2.

5.2 Other Clustering Methods

In our main analysis, we group occupations to levels of hierarchy using a k-means algorithm based on the estimated hierarchy score. A potential drawback of this approach is that it takes into account the overall level of uncertainty in the estimation of ranks to determine the number of hierarchy levels, but does not take into account the uncertainty around each individual rank. In this section, we show that our results are robust to an alternative clustering method that is based on the pairwise statistical difference between adjacent ranks. In particular, occupations that have statistically indistinguishable estimated ranks are subsumed into the same level of the hierarchy. To do so, we order occupations according to their rank and compute rank differences between adjacent pairs. Starting from the highest rank, occupations are grouped together if the rank difference is lower than the estimated standard deviation of the higher ranked occupation. We iteratively apply the grouping procedure until all adjacent occupation pairs are evaluated.

Figure 8b shows that this ad-hoc clustering method leads to a coarser classification of hierarchy levels compared to our baseline. The maximum number of hierarchy levels is 9 instead of 19. Importantly, however, our empirical findings are largely unaffected. Panel D of Table 6 shows that the qualitative conclusion remains unchanged, and the quantitative results are similar. The magnitude of coefficients roughly doubles, which is as expected as the number of hierarchy levels is cut in half.

Table 6: Summary Regressions Hierarchies

	(1) Employment	(2) Number of Occupations	(3) Share External Hires	(4) Share Management	(5) Log Wage	(6) AKM Person Effect
Panel A: Baseline						
Hierarchy Level	-46.95***	0.213***	-0.0849***	0.0441***	0.0429***	0.0400***
	(9.657)	(0.0326)	(0.00542)	(0.00464)	(0.00317)	(0.00295)
Squared Hierarchy Level	2.309***	-0.0195***	0.00385***	-0.000984***	-0.00157***	-0.00148***
	(0.542)	(0.00329)	(0.000430)	(0.000377)	(0.000239)	(0.000220)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Firm Fixed Effects	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.677	0.493	0.301	0.268	0.755	0.655
Worker-year Observations	1,928,885	1,928,885	1,849,341	1,928,885	1,928,885	1,928,885
Panel B: Sample Firm						
Hierarchy Level	-13.71	0.0797	-0.00801	0.220	0.0311	0.0373
Squared Hierarchy Level	0.986	-0.0319	-0.000381	-0.0156	0.00261	0.00206
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.218	0.784	0.135	0.165	0.581	0.569
Worker-year Observations	2,412	2,412	2,412	2,412	2,412	2,412
Panel C: Unadjusted Components						
Hierarchy Level	-41.70***	0.197***	-0.0801***	0.0416***	0.0407***	0.0381***
	(9.689)	(0.0253)	(0.00493)	(0.00432)	(0.00319)	(0.00296)
Squared Hierarchy Level	1.917***	-0.0170***	0.00346***	-0.000963***	-0.00146***	-0.00138***
	(0.560)	(0.00212)	(0.000387)	(0.000355)	(0.000241)	(0.000222)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Firm Fixed Effects	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.674	0.460	0.300	0.258	0.746	0.638
Worker-year Observations	1,978,550	1,978,550	1,893,671	1,978,550	1,978,550	1,978,550
Panel D: Ad-hoc Clustering						
Hierarchy Level	-65.88	2.417***	-0.193***	0.0643***	0.0810***	0.0770***
	(46.82)	(0.730)	(0.0223)	(0.0191)	(0.00763)	(0.00747)
Squared Hierarchy Level	5.462	-0.379***	0.0159***	0.00272	-0.00477***	-0.00493***
	(10.18)	(0.113)	(0.00447)	(0.00347)	(0.00115)	(0.00124)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Firm Fixed Effects	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.854	0.872	0.355	0.388	0.812	0.744
Worker-year Observations	1,928,885	1,928,885	1,877,733	1,928,885	1,928,885	1,928,885

Note: This table shows regression results for regressing various firm and worker characteristics on a quadratic polynomial in the hierarchy level. All regressions include year fixed effects and firm fixed effects. Standard errors are clustered on the firm level and displayed in parentheses.

5.3 Stability over Time

Finally, we turn to the stability of the organizational structure over time. Throughout the paper, we have treated the number of hierarchy levels within a firm as a fixed characteristic. However, while the job ladder is identified using observed worker transitions between occupations in any two years of the observation period, this does not necessarily imply that the firm employs workers in the same set of occupations in every year. We examine the stability of the hierarchical structure by tracking changes in the number of actually populated hierarchy levels over time. Firms can add levels to their job ladders over time, e.g., by opening a new production line, or they can reduce the number of levels, e.g., due to outsourcing. Because firms with only one level mechanically have a fixed occupational structure over time, we exclude them from the sample. Our main analysis of structural stability focuses on the remaining 2,022 firms with at least 2 levels.

In our sample, 55 percent of the firms with more than one level in their hierarchy change their number of levels over time. Among those that change, 35 percent expand by adding one or more levels, 19 percent contract and remove levels, and 46 percent both add and remove levels. Table 7 shows annual employment growth rates for the different types of firms. Most remarkably, employment growth vastly differs between expanding and contracting firms. While 50 percent of firms increase their employment between 2007 and 2014, 71 percent of expanding firms and 28 percent of contracting firms grow. The average firm grows by 3 percent annually.

In our sample, firms with stable hierarchies employ workers at every level in each year and grow slightly over the 2007 - 2014 period. Employment increases by roughly 10 percent per year in the average firm and the annual growth rate of the median firm is 2 percent, see Table 7.

Table 7: Firm growth form 2007 to 2014

	Median Number of Workers in 2007	Share of Growing Firms	Average Annual Growth Rate	Median Annual Growth Rate	Number of Observa-tions
Stable Hierarchies	53	0.44	0.08	0.02	892
Expanding Firms	28	0.71	0.64	0.14	396
Contracting Firms	52	0.28	-0.01	-0.03	218
Both Expanding and Contracting	28	0.53	0.17	0.04	516
All Firms	41	0.50	0.20	0.03	2022

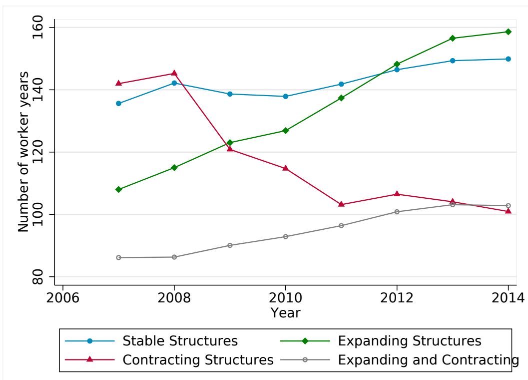
Notes: This table shows average annual employment growth rates of firms with more than one hierarchy level between 2007 and 2014. The sample includes 892 firms with stable structures, 396 expanding firms, 218 contracting firms, and 516 firms that are both expanding and contracting (see further details of the sample restrictions in Table 3)..

Roughly half of the firms in our sample, however, change their hierarchical structure over time. We would thus like to know which firms are expanding or contracting their hierarchical structure. Figure 9 shows firm characteristics of four types of firms over time: Stable firms, which do not change their hierarchy structure, expanding firms, contracting firms, which add or drop levels, and firms with volatile structures that both add and drop levels over time. Firms with stable structures are among the largest firms both in terms of employment and in terms of occupations and they grow at a low rate, see Figures 9a and 9b.

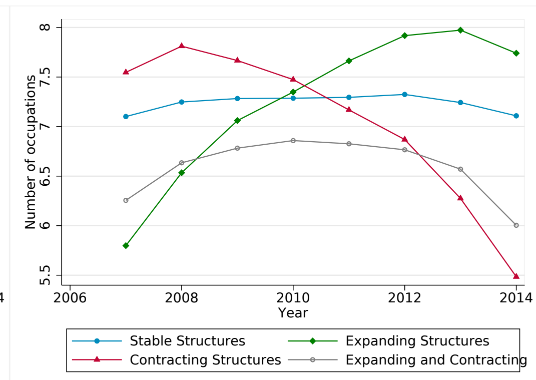
In contrast, expanding firms are strongly growing over time, adding roughly 30 percent of workers and 2 occupations over the 8-year period, while firms with a contracting structure are shrinking in employment and occupations. Firms with volatile structures are smaller than the other firms, with fewest employees on average and fewest occupations. Our finding that in medium-sized firms, strong employment growth is associated with an expansion of the hierarchy structure, is consistent with Caliendo *et al.* (2015a) who show that French manufacturing firms grow by adding additional levels to their hierarchies (see also Caliendo *et al.*, 2015b and Friedrich, 2015). This finding suggests an interesting avenue for future work: The algorithm could be adapted to changes in the firm structure, and be estimated separately before and after a structural change in the organization.

Figure 9: Organizational Structures

(a) Employment



(b) Occupations



Notes: The figures show the development of employment, number of occupations, and log wages over time in firms with different types of structures. The sample includes 892 firms with stable structures, 396 expanding firms, 218 contracting firms, and 516 firms that are both expanding and contracting (see further details of the sample restrictions in Table 3).

6 Conclusion

In this paper, we have developed a method to study how promotion dynamics and wages are shaped by the structure of internal job hierarchies in firms. Our paper addressed the conceptual challenge related to ordering different job titles in firms into levels of a hierarchy by using observed flows between jobs *within* firms. In the first step of our approach, we identified networks of occupations based on observed occupational transitions within the firm. These networks were allowed to be segmented, so that a firm can consist of multiple internal labor markets. We built on the approach in Kline *et al.* (2019) to remove workers that break an internal market into further submarkets. In the second step, we extended the hand-curated approach of Baker *et al.* (1994), and established a data-driven ranking of occupations within the internal labor market by exploiting the direction of flow frequencies between occupations. Finally, we grouped occupations by k-means clustering and chose the optimal number of levels following Bonhomme *et al.* (2019).

We applied our method on linked employer-employee data from Norway that reports fine-grade occupational codes and tracks contract changes within firms. Our evidence from 2000 large firms can be summarized with three main findings. First, we documented that the structure of internal labor markets differs widely across firms. Second, we offered a wide range of empirical evidence that supported theories of careers in organizations. Finally, we found strong evidence that the wage setting process in the average firms appears to be partly determined by systems of promotions. These findings are in line with tournament theory that predicts an increase in the wage spread between hierarchy levels toward the top of the job ladder. Yet, we also found that average person fixed effects from AKM wage regressions increase at higher hierarchy levels. This evidence lends some support to models where firms use the hierarchy to allocate more talented individuals towards higher levels.

In this paper, we scraped the surface of how wage setting in firms depends on internal labor markets. Yet, our method has the potential to advance our understanding of wage policies in firms and how firms affect the wage distribition (e.g., Card *et al.*, 2013; Song *et al.*, 2018; Lamadon *et al.*, 2019). While most studies can only speculate about the mechanisms that underlie the increasing dispersion in firm quality and the rise in sorting of high-wage workers to high-wage firms, our framework allows for an assessment of the role of internal labor markets in shaping the wage policy of firms. This opens multiple avenues for future research, e.g., it enables a comprehensive assessment of how internal labor markets of firms affect the decision to outsource, related to Goldschmidt & Schmieder (2017) who documented impacts of outsourcing certain occupations, such as cleaners or security personnel, on the wage structure in Germany. Moreover, our framework could be used in future research to study whether workers are willing to take wage cuts in order to access internal labor markets with more promising wage prospects.

References

ABOWD, JOHN M., KRAMARZ, FRANCIS, & MARGOLIS, DAVID N. 1999. High Wage Workers and High Wage Firms. *Econometrica*, **67**(2), 251–333.

ABOWD, JOHN M., CREECY, ROBERT H., & KRAMARZ, FRANCIS. 2002. *Computing Person and Firm Fixed Effects Using Linked Longitudinal Employer-Employee Data*. Working Paper TP-2002-06. US Census Bureau: Longitudinal Employer Household Dynamics.

ANDREWS, M. J., GILL, L., SCHANK, T., & UPWARD, R. 2008. High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **171**(3), 673–697.

BAKER, GEORGE, GIBBS, MICHAEL, & HOLMSTROM, BENGT. 1994. The internal economics of the firm: Evidence from personnel data. *The Quarterly Journal of Economics*, **109**(4), 881–919.

BARTH, ERLING, MOENE, KARL O, & WILLUMSEN, FREDRIK. 2014. The Scandinavian model—an interpretation. *Journal of Public Economics*, **117**, 60–72.

BAYER, CHRISTIAN, & KUHN, MORITZ. 2019. Which Ladder to Climb? Decomposing Life Cycle Wage Dynamics.

BERTRAND, MARIANNE, & SCHOAR, ANTOINETTE. 2003. Managing with style: The effect of managers on firm policies. *The Quarterly journal of economics*, **118**(4), 1169–1208.

BERTRAND, MARIANNE, BURGESS, ROBIN, CHAWLA, ARUNISH, & XU, GUO. 2020. The glittering prizes: Career incentives and bureaucrat performance. *The Review of Economic Studies*, **87**(2), 626–655.

BOGNANNO, M.L. 2001. Corporate Tournaments. *Journal of Labor Economics*, **19**, 290–315.

BONHOMME, STEPHANE, LAMADON, THIBAUT, & MANRESA, ELENA. 2019. *Discretizing Unobserved Heterogeneity*. Working Paper.

CALIENDO, LORENZO, MONTE, FERDINANDO, & ROSSI-HANSBERG, ESTEBAN. 2015a. The anatomy of French production hierarchies. *Journal of Political Economy*, **123**(4), 809–852.

CALIENDO, LORENZO, MION, GIORDANO, OPROMOLLA, LUCA DAVID, & ROSSI-HANSBERG, ESTEBAN. 2015b. *Productivity and organization in Portuguese firms*. Tech. rept. National Bureau of Economic Research.

CARD, DAVID, HEINING, JOERG, & KLINE, PATRICK. 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*, **128**(3), 967–1015.

CESTONE, GIACINTA, FUMAGALLI, CHIARA, KRAMARZ, FRANCIS, & PICA, GIOVANNI. 2019. *Insurance Between Firms: The Role of Internal Labor Markets*. Working Paper 11336. CEPR.

CHIAPPORI PIERRE, ANDRE, & SALANIE, BERNARD. 2003. Testing Contract Theory: A Survey of Some Recent Work. *Advances in Economics and Econometrics*, **1**.

CLAUSET, AARON, ARBESMAN, SAMUEL, & LARREMORE, DANIEL B. 2015. Systematic inequality and hierarchy in faculty hiring networks. *Science Advances*, **1**(1).

DEVARO, JED, & KAUHANEN, ANTTI. 2016. An Opposing Responses Test of Classic Versus Market-Based Promotion Tournaments. *Journal of Labor Economics*, **34**(3), 747–779.

DOERINGER, PETER, & PIORE, MICHAEL. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, MA: D. C. Heath and Company (Armonk, NY: London: M. E. Sharpe, Inc., reprinted 1985).

ERIKSSON, TOR. 1999. Executive Compensation and Tournament Theory: Empirical Tests on Danish Data. *Journal of Labor Economics*, **17**(2), 262–280.

FRIEDRICH, BENJAMIN. 2015. *Trade shocks, firm hierarchies and wage inequality*. University of Aarhus, Department of Economics.

GABAIX, XAVIER, & LANDIER, AUGUSTIN. 2008. Why has CEO pay increased so much? *The Quarterly Journal of Economics*, **123**(1), 49–100.

GARICANO, LUIS, & HUBBARD, THOMAS N. 2007. Managerial leverage is limited by the extent of the market: Hierarchies, specialization, and the utilization of lawyers' human capital. *The Journal of Law and Economics*, **50**(1), 1–43.

GARICANO, LUIS, & ROSSI-HANSBERG, ESTEBAN. 2006. Organization and inequality in a knowledge economy. *The Quarterly Journal of Economics*, **121**(4), 1383–1435.

GIBBONS, ROBERT, & WALDMAN, MICHAEL. 1999a. Careers in organizations: Theory and evidence. *Handbook of labor economics*, **3**, 2373–2437.

GIBBONS, ROBERT, & WALDMAN, MICHAEL. 1999b. A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics*, **114**(4), 1321–1358.

GIROUD, XAVIER, & MUELLER, HOLGER M. 2019. Firms? Internal Networks and Local Economic Shocks. *The American Economic Review*, forthcoming.

GOLDSCHMIDT, DEBORAH, & SCHMIEDER, JOHANNES F. 2017. The rise of domestic outsourcing and the evolution of the German wage structure. *The Quarterly Journal of Economics*, **132**(3), 1165–1217.

KLINE, PATRICK, SØLVSTEN, MIKKEL, & SAGGIO, RAFFAELE. 2019. *Leave Out Estimation of Variance Components*. Working Paper.

KRUEGER, ALAN B., & SUMMERS, LAWRENCE H. 1988. Efficiency Wages and the Inter-Industry Wage Structure. *Econometrica*, **56**(2), 259–293.

LAMADON, THIBAUT, MOGSTAD, MAGNE, & SETZLER, BRADLEY. 2019. *Imperfect Competition, Compensating Differentials, and Rent Sharing in the U.S. Labor Market*. Working Paper 25954. NBER.

LAZEAR, EDWARD P., & OYER, PAUL. 2004. The Structure of Wages and Internal Mobility. *The American Economic Review*, **94**(2), 212–216.

LAZEAR, EDWARD P., & ROSEN, SHERWIN. 1981. Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, **89**(5), 841–864.

LAZEAR, EDWARD P, & SHAW, KATHRYN L. 2007. Personnel economics: The economist's view of human resources. *Journal of economic perspectives*, **21**(4), 91–114.

LAZEAR, EDWARD P, & SHAW, KATHRYN L. 2009. Wage structure, raises and mobility: An introduction to international comparisons of the structure of wages within and across firms. *Pages 1–57 of: The structure of wages: An international comparison*. University of Chicago Press.

LAZEAR, EDWARD P, SHAW, KATHRYN L, & STANTON, CHRISTOPHER T. 2015. The value of bosses. *Journal of Labor Economics*, **33**(4), 823–861.

MCCUE, KRISTIN. 1996. Promotions and wage growth. *Journal of Labor Economics*, **14**(2), 175–209.

NEAL, DEREK. 1999. The Complexity of Job Mobility among Young Men. *Journal of Labor Economics*, **17**(2), 237–261.

SATTINGER, MICHAEL. 1975. Comparative advantage and the distributions of earnings and abilities. *Econometrica: Journal of the Econometric Society*, 455–468.

SCHUBERT, GREGOR, STANSBURY, ANNA, & TASKA, BLEDI. 2020. *Monopsony and Outside Options*. Working Paper.

SONG, JAE, PRICE, DAVID J, GUVENEN, FATIH, BLOOM, NICHOLAS, & VON WACHTER, TILL. 2018. Firming up inequality. *The Quarterly journal of economics*, **134**(1), 1–50.

TOPEL, ROBERT H., & WARD, MICHAEL P. 1992. Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics*, **107**(2), 439–479.

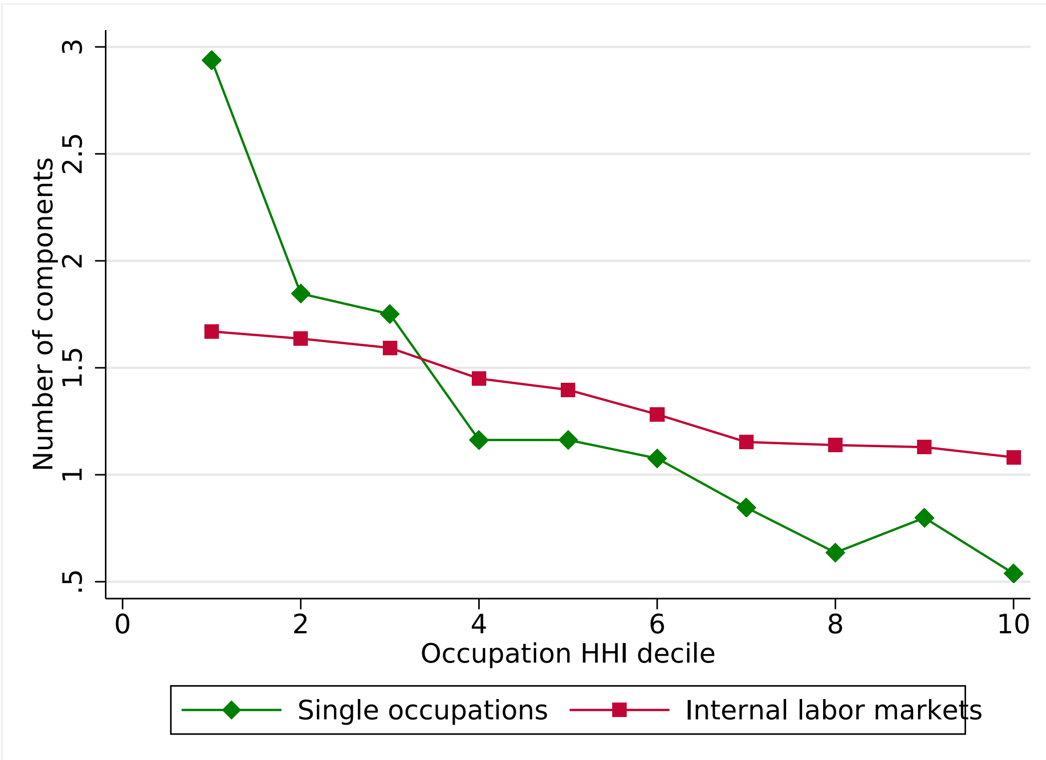
VAN DER KLAAUW, BAS, & DA SILVA, ANTÓNIO DIAS. 2011. Wage dynamics and promotions inside and between firms. *Journal of Population Economics*, **24**(4), 1513–1548.

WALDMAN, MICHAEL. 1984. Job Assignment, Signalling, and Efficiency. *RAND Journal of Economics*, **8**, 255–287.

WANG, H, & SONG, M. 2011. Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. *The R Journal*, **3**(2), 29–33.

A Appendix: Additional Tables and Figures

Figure A1: Internal Labor Markets and Occupational Concentration



B Details on the Algorithms

B.1 Internal Labor Markets

We identify ILMs as connected sets in the network of occupations within firms. Let $G = (U, V, E)$ denote a firm-specific bi-partite graph where U denotes the set of workers in the firm, V denotes the set of occupations, and E denotes the set of links between workers and occupations. A connected component is a subgraph of G in which any two occupations are connected to each other by some path of worker transitions, and which is connected to no additional occupations in G. Computing connected components is a prerequisite for several applications in Economics, e.g., the wage decomposition in AKM. Functions that identify components in a network typically rely on a breadth-first search algorithm and are available in standard statistical packages such as *igraph* in R or *Boost Graph Library* in MATLAB.

In order to account for potential measurement error in occupational coding, we propose a leave-X-percent-out procedure that prunes the data. This pruning procedure requires that the bi-partite network remains connected when any one worker is removed. This boils down to finding workers that constitute cut vertices or *articulation points* in the bi-partite network. In contrast to the standard procedure of identifying these cut points, we cut the network only if removing a single worker destroys less than X percent of the vertices that are attached to an occupation.

Algorithm 1: Pruning Network

Result: Leave-X-percent-out Component Structure

compute degree d_v for each occupation in $v \in V$;

construct G' where each link that enters occupation $v \in V$ is duplicated $100/(d_v * X)$ times ;

construct G'_1 from G' by deleting all workers that are articulation points in G' ;

remove duplicated links and return G_1 ;

B.2 Minimum Violations Ranking

We rank occupations within each ILM using a minimum violations ranking (Clauset *et al.*, 2015). In contrast to the first step of our algorithm, the direction of the worker transition that links two occupations is important for this step. Let $H = (V, T)$ denote a network of occupations within a firm V that are connected by directed and weighted worker transitions, T. Our goal is to define a ranking of all occupations $v \in V$ such that the number of "violations" is minimized. A violation is a worker transition (u, v) where the rank of the origin occupation u is higher than the rank of the target occupation v. Complex networks like job to job transitions within firms often exhibit multiple minimum violation rankings, in which several distinct orderings of occupations produce the same smallest number of violations. We apply the sampling procedure proposed in Clauset *et al.* (2015) to find a consensus ranking in the case of several possible optimal rankings. The algorithm starts from an initial ranking that sorts occupations according to their out-degree. At each step in the algorithm, a randomly chosen pair of occupations is chosen and a ranking is proposed in which their ranks are swapped. We accept each proposal that leads to a lower or equal number of violations. Once the algorithm has converged on the minimum number of violations, this might still lead to different accepted rankings (since we are also accepting neutral proposals). We then sample from the set of rankings with the minimum number of violations and average the ranks for each occupation to get a consensus ranking. We repeat the procedure R times to compute the standard deviation of mean ranks across repetitions as our measure of uncertainty.

Algorithm 2: Minimum Violations Ranking Algorithm

Result: Average mean ranking over R repetitions, std. dev. of mean rankings

for i = *1* **to** R **do**

  sort occupations according to out-degree (decreasing);

  compute number of violating links S ;

  set $t = 0$;

  while $t < T$ **do**

    $t = t + 1$;

    switch ranks of two randomly chosen occupations ;

    compute number of violating links S' ;

    if $S' \leq S$ **then**

      if $S' < S$ **then**

        $S = S'$;

        delete previously stored rankings ;

        $t = 1$;

      end

      store updated ranking ;

    end

  end

  store number of violations and average ranking over samples ;

end

compute summary stats over all repetitions: mean ranking, std. ranking ;

B.3 Clustering Ranks to Hierarchy Levels

In order to group occupations into hierarchy levels, we apply a one-dimensional kmeans clustering algorithm based on the estimated average rank r_v for each occupation $v \in V$. In particular, we partition the N occupations within a firm into$K$ groups, corresponding to group indicators $\hat{k}_v \in 1, \ldots, K$ in order to minimize the objective function

$$(\hat{h}, \hat{k}_1, \ldots, \hat{k}_N) = \arg \min_{h, k_1, \ldots, k_N} \sum_{v=1}^{N} (r_v - h(k_v))^2 \tag{B1}$$

where $\hat{h}(k)$ is the mean of r_v in group $\hat{k}_i = k$ and K is a pre-determined number of hierarchy levels. Wang & Song (2011) provide an optimal dynamic programming algorithm to minimize (B1) in this one-dimensional setting that leads to guaranteed optimality and is reproducible since it does not rely on computational approximations such as kmeans in higher dimensions.

Following Bonhomme *et al.* (2019), we choose the number of hierarchy levels, K, such that the within-group variation in ranks is of the same magnitude as the overall noise level in the estimation of the ranks. Specifically, let $\hat{Q}(K) = \frac{1}{N} \sum_{v=1}^{N} (r_v - \hat{h}(k_v))^2$ denote the value of the kmeans objective function corresponding to K groups. Then, we choose

$$\hat{K} = \min_{K \geq 1} \left\{ K : \hat{Q}(K) \leq \frac{1}{N-1} \sum_{v=1}^{N} Var(r_v) \right\} \tag{B2}$$

where $Var(r_v)$ is the variance of the estimated rank obtained from the minimum violations ranking algorithm.