Identifying Career Patterns in Online Labor Markets: A Sequence and Cluster Analysis

Mareike Seifried
LMU Munich
m.seifried@lmu.de

Abstract

Career research resolves around the question how workers progress through their working life (cf. Ng, Eby, Sorensen, & Feldman, 2005) and the concept of a career has usually been studied against the backdrop of an organization structured by internal labor markets (O’Mahony & Bechky, 2006). However, today, it is estimated that up to 162 million individuals already work as independent workers in external labor markets and it is predicted that their share will grow very rapidly in the future (Mayinka et al., 2016). As part of this trend, an increasing number of gig workers participates in online labor markets (OLMs) and works on a completely digital basis. These steadily growing OLMs refer to markets “where labor is exchanged for money, the product of that labor is delivered over a wire and the allocation of labor and money is determined by a collection of buyers and sellers operating within a price system” (Horton, 2010: 516). Today, OLMs facilitate the allocation of productive effort across global economies and are seen as an important new labor pool for a variety of tasks (Malone et al., 2010), ranging from micro-tasks in translations and transcriptions on Amazon Mechanical Turk to high-skill software development projects on UpWork.

Due to the changing nature of employment relationships, career researchers have begun to question the relevance of an organizational framework as a backdrop to career progression. By being confronted with multiple career choices and lacking an organizational framework that helps to link those choices in a progressive manner, gig workers have to act as independent creators of their own careers. As a result, scholars have started to explore how gig workers try to achieve career progression without the benefit of organizational guidance and find that careers become less predictable but much more dynamic and multi-directional than traditional corporate careers (O’Mahony & Bechky, 2006; Reilly, 2017).

However, existing research on careers in external labor markets has so far neglected the increasing digitization of contract work (cf. Agrawal et al., 2016). That is, the trend towards independent work was to a large extent enabled by and is further fueled by technological advancements, enabling workers to provide their services through digital matching platforms and changing not only the nature of work but also how careers might unfold. Given the increasing relevance of these markets, the heterogeneity of the digital workforce and outsourced tasks as well as the volatility of this work environment, it is important to examine how digital workers use OLMs and self-assemble their careers. Thus, we ask:
Which career paths emerge in online labor markets for knowledge tasks? We contribute to existing research on careers in external markets by studying whether and how careers unfold in this novel and purely digital context. Career paths refer to models or prototypes describing the career sequences of a cluster of individuals (Joseph, Fong, Ang, Slaughter, 2012). Prior work on career paths has analyzed career sequences of musicians (Abbott and Hrycak, 1990), women in finance occupations (Blair-Loy, 1999), and IT professionals (Joseph et al., 2012). However, OLMs represent global labor pools for a variety of tasks and online freelancing differs from offline freelancing in several ways. First, digital workers face even more career choices, i.e. the flexibility to work on heterogeneous tasks and for different employers from different countries at the same and usually only for a short amount of time. Thus, they are exposed to an even larger pool of potential employers, task types and work settings, making it even more challenging to structure their careers. Consequently, some even question the existence of careers in these spot markets for tasks because participants might only work on a loosely coupled number of jobs with no pre-defined path for steady upward movement (cf. Chen & Horton, 2016). Second, career choices and performance are transparent. Workers create a profile that documents their working experience on the platform as well as the feedback given by past employers. Due to the unified reputation system, performance is easily comparable between workers and task categories. On the one hand, this provides the opportunity to observe behavioral patterns over time and compare progression in important career outcomes across different work settings. On the other hand, it might also affect what freelancers focus on because future employers can observe their actions and performance. Third, the role of social capital for career progression is almost eliminated. In the offline context, scholars have emphasized the importance of social ties and mentorships for career progression (Seibert, Kraimer, Liden, 2001; Reilly, 2017). In contrast, progression in OLMs can be fully attributed to an individual’s actions. Finally, existing studies on careers in the gig economy have primarily focused on full-time gig workers in highly specific contexts (e.g. stand-up comedians). However, the majority of digital workers does not work full-time on the platform. Relatedly, the gig workforce in OLMs is very diverse in motivations, demographics, origin, and educational background (Mayinka et al., 2016). Consequently, self-assembled careers in OLMs might unfold in diverging patterns and do not necessarily follow the same overall logic or path.

To develop a taxonomy of careers in OLMs, i.e. understanding whether and how careers unfold, we aim at identifying and analyzing differential patterns of career progression and activities (e.g. career length, engagement level, task specialization vs. diversification) and identifying the characteristics of individuals associated with a given career pattern. We adopt an inductive and quantitative approach to examine the career paths of individuals working on the platform. Identifying career sequences requires analyzing work histories across individuals and over time. We conduct an optimal matching (a sequence analysis technique) and cluster analyses using an extensive dataset from Upwork.com, the world’s largest freelancing website. The panel dataset was crawled in 2017 and consists of over 60,000 freelancers and more than 2 million jobs posted on the platform and includes very detailed information on freelancers, employers, work histories, and job characteristics.
Identifying Career Patterns in Online Labor Markets:

A Sequence and Cluster Analysis

Submitted by Mareike Seifried

*Ludwig-Maximilian-University (LMU) Munich*

Working Paper (December 2018)

PRELIMINARY AND INCOMPLETE
1. Introduction

Career research resolves around the question how workers progress through their working life (Ng, Eby, Sorensen, & Feldman, 2005) and the concept of a career has usually been studied against the backdrop of an organization structured by internal labor markets (O’Mahony & Bechky, 2006). A career was traditionally symbolized by the “organization man” (Whyte, 1956) climbing a career ladder comprising of orderly sequences of jobs with increasing responsibility, income and status within a single firm (Hall, 1976). But the advent of the gig economy threatens the traditional career model because a growing number of workers “is no longer employed in jobs with a long-term connection with a company but are hired for gigs under flexible arrangements as independent contractors or consultants, working only to complete a particular task or for defined time and with no more connection with their employer (…)”. (Friedman, 2014: 171)

In fact, today, it is estimated that up to 162 million individuals already work as independent workers in external labor markets and it is predicted that their share will grow very rapidly in the future (Mayinka et al., 2016). Opposed to individuals employed in an organization, gig workers have to act as independent creators of their own careers. They are confronted with multiple career choices and lack an organizational framework that helps to link those choices in a progressive manner.

Due to the changing nature of employment relationships, career researchers have begun to question the relevance of an organizational framework as a backdrop to career progression. Since workers have multiple options, there is no single way for reaching success and career paths become multidirectional (Baruch, 2003). As a result, scholars have started to explore career development of gig workers and qualitative evidence suggests that careers in external labor markets are less predictable but much more dynamic and multi-directional than traditional corporate careers (O’Mahony & Bechky, 2006; Reilly, 2017).

Although existing research on careers in external labor markets has improved our understanding of non-traditional career patterns, we lack quantitative studies analyzing career progression. At the same time, scholars have so far neglected the increasing digitization of contract work (cf. Agrawal et al., 2016). However, the trend towards independent work was to a large extent enabled by and is further fueled by technological advancements, enabling workers to provide their services through digital matching platforms. Particularly and as part of this trend towards platform-enabled work, an increasing number of gig workers participates in online labor markets (OLMs) and works on a completely digital basis. These steadily growing OLMs refer to markets “where labor is exchanged for money, the product of that labor is delivered over a wire and the
allocation of labor and money is determined by a collection of buyers and sellers operating within a price system” (Horton, 2010: 516). Today, OLMs facilitate the allocation of productive effort across global economies and are seen as an important new labor pool for a variety of tasks (Malone et al., 2010), ranging from micro-tasks in translations and transcriptions on Amazon Mechanical Turk (Mturk) to large-scale innovation projects on Innocentive.

Given the volatility and digital nature of this environment, the short-term employment relationships and the heterogeneity of workers, employers, tasks, it is unclear how careers might unfold. Specifically, due to the volatility and flexibility of OLMs, digital workers face a variety of career choices: In OLMs, workers have to decide on which task (job), in which task category (occupation), for which employer (organization) from which country (geography), for which price and contract (contractual terms) they want to work and whether the job should be done via an agency (affiliation). Thus, the question arises whether there are distinct patterned trajectories in careers or whether careers of digital workers are random and accidental. Since OLMs are predicted to grow fast in the future, it is important to examine how digital workers self-assemble their careers to achieve progression in important career outcomes. Specifically, we are interested in developing a taxonomy of career patterns in OLMs, i.e. understanding the differential types of careers that might emerge.

As a first step, we analyze whether there exist typical patterns in career mobility or more specifically wage mobility, i.e. the sequence of upward, lateral, downward movement in the pay distribution. By examining progress in terms of hourly wages, we can answer questions about the direction of careers and whether there exist differences across freelancers. For example, some freelancers might be able to enter at high pay ranks and pursue a relatively stable career. Others might show a steady upward movement similar to organizational career ladders or enter at low pay ranks and never make it to the top of the pay distribution. This is based on the assumption that a true career only exists if there is some sort of progression (Rosenfeld, 1992).

Second, we aim at analyzing the underlying career moves and job shifts that structure these patterns, e.g. moves between task categories, employers, and countries. Put differently: How are different ways to the top of the pay hierarchy generated? Third, we identify and describe the individual profiles of freelancers following these different paths in terms of demographic characteristics. This helps to understand which types of freelancer are actually able to build “online careers”.

3
Identifying career paths requires analyzing work histories across individuals and over time. We adopt an inductive and quantitative approach and use an optimal matching algorithm (a sequence analysis technique) and cluster analysis. This approach is well-established in career research and has been applied to a variety of contexts, e.g. careers of musicians (Abbott & Hryck, 1990), women in finance occupations (Blair-Loy, 1999), and IT professionals (Joseph, Fong Boh, Ang, Slaughther, 2012).

We use an extensive and novel dataset from Upwork.com, the world’s largest freelancing website. The panel dataset was crawled in 2017 and consists of over 60,000 freelancers and more than 2 million jobs posted on the platform and includes detailed information on freelancers, employers, work histories, and job characteristics.

We first review career theory and existing work on careers in external labor markets. We then define online labor markets and describe the empirical context of our analysis, the platform Upwork. After applying career theory to our specific context, we describe our data, the applied method and current status of our analysis.

2. Career Theory and Prior Work
2.1 Career Theory
Career theory postulates that workers move across jobs in a structured progression (Rosenfeld 1992, Schein 1978, Spilerman 1977). In general, career theory builds on a combination of perspectives from psychology, sociology, anthropology, and economics rather than referring to a unified theory (O’Mahony & Bechky, 2006). Thus, all attempts to explain the unfolding of human work experience over time can be considered career theories (Arthur et al., 1989).

A career is generally defined as a sequence of jobs or work-related experiences within an individual’s work history (Spilerman, 1977) that “form a unique pattern over the individual’s life span” (Sullivan & Baruch, 2009: 1543). Job shifts or transitions to different jobs, occupations or firms can be seen as the “building blocks” of careers (Rosenfeld, 1992; Sullivan & Baruch, 2009), often referred to as job mobility (Feldman & Ng, 2007; Sicherman & Galor, 1990; Rosenfeld, 1992). Hence, career researchers usually look for underlying job shifts that structure careers and thus form patterns in job histories, although "random" careers are possible (Rosenfeld, 1992). Career paths or patterns (we will use these terms synonymously) then refer to models or prototypes describing the career sequences of a cluster of individuals (Joseph, Fong, Ang, Slaughther, 2012).
The term career is generally associated with some sort of progression over time or at least coherence regarding the jobs a person holds over the work life (Wilensky, 1960). That is because traditional career models had a clear, unidimensional or linear direction of "advancement" (Rosenbaum, 1979; Wilenky, 1964): The organizational hierarchy was perceived as a career ladder to climb up. *Career success* is usually evaluated via the rate of upward mobility (promotion) and external indicators of achievement (e.g. salary and social status) (Rosenfeld, 1992).

2.2 Career Development in Internal Labor Markets

In internal labor markets, career ladders were seen as stable and clearly defined paths to the top, implying clear prescriptions and expectations on what to do to climb up (Baruch, 2003). Internal labor markets allocate labor, determine wages, and define the scope of specific jobs according to sets of rules and procedures (Doeringer & Piore, 1971). Employees are rewarded with increasing salaries and promotions along job ladders in return for their commitment and long-term employment (Baker, Gibbs, & Holmstrom, 1994; Grandjean, 1981; Osterman, 1984). These structures help foster a shared understanding of how one job follows the next and the associated skill requirements (Lawrence, 1990).” (O’Mahony & Bechky, 2006)

However, recent economic, technological, and social changes have altered traditional employer-employee relationships and the work context, creating changes in how individuals enact their career (Sullivan et al., 2009), both in internal and external labor markets. As a consequence, a literature on boundaryless careers has emerged (e.g., Arthur & Rousseau, 1996). Boundaryless careers do not unfold in a single organization but comprise less structured sequences of jobs that may cross occupational, organizational, and geographical boundaries (Arthur & Rousseau, 1996). Career research has thus begun to address the increasing mobility of workers and its consequences for career development, for example by introducing concepts such as interorganizational career ladders (Bidwell & Briscoe, 2010). Nevertheless, they are unable to capture the reality of workers developing a career in external labor markets, both online and offline.

2.3 Career Development in External Labor Markets

Faced with numerous career choices and in the absence of career ladders to sequence jobs, workers in external labor markets must develop strategies to craft and take responsibility for
their own careers. At the same time, the volatility of the work environment makes career progression much less predictable (O’Mahony & Bechky, 2006).

Existing studies are usually interested in why and how individuals move into and out of self-employment (e.g. Carroll & Mosakowski, 1987) but rarely how self-employed workers progress through their careers. One of the few exceptions is a study conducted by O’Mahony and Bechky (2006), who examine how independent (high-technology and film production) workers try to achieve career progression. They analyze how workers manage the so-called “career progression paradox”, i.e. the problem of finding a job to develop one’s skills and employability without prior work experience related to these skills. Thus, they focus on job transitions events and develop the concept of “stretchwork”, referring to jobs bridging from proven competencies to new ones to help reconcile the career paradox.

Prior work on market-based careers has also proposed that social networks might structure external labor markets (Faulkner & Anderson, 1987). Analyzing careers in the film industry, Faulkner and Anderson (1987) find that career opportunities differ for those in the core versus the periphery of the industry’s external labor market network. Similarly emphasizing the role of social networks, Reilly (2017) analyzes careers of stand-up comedians and introduces the model of a layered career where each layer consists of a durable social infrastructure. The author suggests that individuals move progressively through three layers but also continuously backwards, suggesting the multidirectional instead of linear nature of freelancing careers.

Although these studies have clearly advanced our understanding of careers in external labor markets, they are qualitative in nature and do not analyze the existence and emergence of differences in complete work histories. At the same time, the context of online labor markets differs from offline freelancing in a number of aspects so that these studies are only partly applicable to this novel context, for example the role of durable social networks. In the following section, we will briefly define and describe online labor markets and how career development might change in this context.

3. Career Development in Online Labor Markets
3.1 Online Labor Markets

Online labor markets (OLMs) such as Amazon Mechanical Turk (MTurk), Upwork (formerly Elance-oDesk), Fiverr, Freelancer.com, Zooniverse, and Innocentive are all platforms that facilitate the allocation of labor across global economies (Agrawal et al., 2015). OLMs are markets “where labor is exchanged for money, the product of that labor is delivered over a wire and
the allocation of labor and money is determined by a collection of buyers and sellers operating within a price system” (Horton, 2010: 516). We focus on spot markets for tasks, a particularly powerful new way of accomplishing work online (Horton, 2010), where employers “buy discrete chunks of labor from a global pool of workers at a market price, similar to how they obtain any other factor of production” (Chen and Horton, 2016: 414).

We focus on Upwork, the world’s largest freelancing website. Upwork was founded in April 2014 following a merger between oDesk and Elance, and was renamed from Elance-oDesk to Upwork in May 2015. Upwork facilitates transactions ranging from administrative support and graphic design to software and web development. Freelancers are earning more than $1bn annually on the site and cover over 3,500 skills (Upwork, 2018). Upwork encourages longer-term projects and encourages high-value, ongoing work (Pofeldt, 2016), for example developing an online marketing strategy (Sales and Marketing), porting an Android app from an iOS app and adding new features to an existing app designed with C++ (both Web, Mobile, and Software Development).

To post projects on Upwork, employers register by providing their contact details and basic information on their firm. Employers can then post any number of jobs and hire as many freelancers as they like, in general and for a single project. Postings include a task description, the employer’s location, the type of contract offered (fixed price or hourly wage), and other job features. For fixed-price projects, employers must specify outcome, budget and deadline. Both parties negotiate a bid for the full project or break it down into milestones. The freelancer submits the deliverable and the employer reviews it, approves the milestone and releases funds after approval. For hourly-wage projects, employers give the expected number of weeks and hours per week to complete the project. Employers can review work through a virtual monitoring system, which tracks billable time and records completed work. The work diary counts keystrokes and takes random screenshots of the freelancer’s screen, enabling the verification of work progress and billable hours.

Workers register by giving their contact details, name and location as well as setting up a profile page. Profile pages include a description of skills, education, work experience outside of the platform, skill test scores, certifications, agency affiliation, portfolio items, and platform work history and feedback scores. Freelancers apply by submitting cover letters and bids to jobs. Employers can interview and negotiate over bids with applicants before hiring. When hiring multiple freelancers for single projects, employers send separate job offers and outline the terms of each contract, which can then deviate from the original job posting.
Once hired, freelancers complete tasks remotely. Submission of deliverables and payments are done via the platform, which then charges a service fee. After completion, employers evaluate the project performance with a feedback score from 1 to 5, on six criteria related to process and outcome quality.

3.4 Differences between Online and Offline (External) Labor Markets

As evident from the description above, OLMs differ from offline labor markets in several ways, which has important implications for career development.

First, work is entirely performed online rather than by physically collocated workers (Chen & Horton, 2016). Given the on-demand and geographically distant nature of online work, workers can work for several employers simultaneously. Hence, OLMs are characterized by many-to-many connections, with some lasting only a few minutes or hours (Felstiner, 2011). The short-term and on-demand nature implies that employers have even lower incentives to train freelancers compared to offline freelancing and training itself might become difficult. Furthermore, freelancers have to continuously search and apply for new jobs and job mobility is extremely high. At the same time, freelancers have to decide on their own what to do next and for which price.

Second, OLMs are characterized by multiple layers of heterogeneity. Most OLMs broker highly heterogeneous tasks, enabling workers to work in diverse task categories and on tasks with different skill levels. Prior work has indeed shown that freelancers use this flexibility and switch between distinct task categories (Kokkodis & Ipeirotis, 2016; Leung, 2014). Employers are flexible in which tasks to outsource and how to specify each job (e.g. contract type, engagement level, skills required, hiring multiple workers). As a consequence, digital workers face even more career choices than freelancers in an offline context. Specifically, they have the flexibility to work on heterogeneous tasks and for different employers from different countries and under different contractual terms. Thus, they are exposed to an even larger pool of potential employers, task types and work settings, making it even more challenging to structure their careers. The low entry barriers and availability of different types of jobs attracts a global digital workforce, which is very diverse in motivations, demographics, origin, and educational background (Mayinka et al., 2016). Consequently, self-assembled careers in OLMs might unfold in diverging patterns and do not necessarily follow the same overall logic or path because free-
lancers have different motivations and experiences. At the same time, a global workforce implies that freelancers from different countries compete for the same jobs although they live in countries with very different living costs and tax regulations.

Third, past career choices are **transparent**, i.e. fully observable by other freelancers and future employers. OLMs offer employers more standardized and complete information on applicants (Agrawal et al., 2016), such as their work history, feedback given by past employers, skills, and ongoing progress of the commissioned work. Importantly, this includes past payments, i.e. workers have transparent pay histories and past wages will inevitably serve as a signal to future employers and anchor for wage negotiations. Further, advanced communication and information technology enable almost real-time access to the worker and her progress. Thus, these choices have an impact on career development.

Forth, the role of **social capital** for career progression is almost eliminated. In the offline context, scholars have emphasized the importance of social ties and mentorships for career progression (Seibert, Kraimer, Liden, 2001; Reilly, 2017). In contrast, progression in OLMs can be fully attributed to an individual’s actions.

Given these differences, it is unclear whether we observe some sort of structured progression and clear career paths or whether all careers are random and accidental. Recent work in the context of OLMs, however, suggests that at least some freelancers have a long-term perspective and strategically apply to jobs and build their work history. Horton and Tamble (2017) find that digital workers are forward-looking and behave strategically in building their skills when demand in a specific skill changes, suggesting that at least some workers actively craft their careers and have a long-term perspective. Leung (2014) studies whether the order of a freelancer’s work history, i.e. the chronological order of job categories the freelancer has worked in, affects employers’ hiring decisions. He finds that employers prefer applicants who move incrementally between similar jobs to those who do not move (specialize in one job category) or those with highly diverse job histories (move between highly dissimilar job categories). These results suggest that career histories matter and that employers favor workers that are committed to a certain job area, but attempt to develop their skills and careers at the same time. Some studies also show that career moves in early stages of an online career are crucial for further career development, e.g. working through an agency to help inexperienced freelancers to get their careers started.
We contribute to this stream by analyzing whether there are common patterns among a set of job history sequences (the pattern question) and if so, how they are produced (the generation question). Specifically, we focus on hourly wages as career success measure to examine career progression and identify wage mobility patterns to examine the direction (upward, lateral, downward, multi-directional) of careers in OLMs. Second, we examine the underlying shifts and career choices of these career patterns (e.g. switching task categories, working for an agency or specific types of employers etc.). We are particularly interested in identifying career paths that lead to the top of the pay distribution (i.e. becoming a top earner in this context). Third, we aim at describing the individual profiles associated with these patterns (e.g. education, gender, origin, platform tenure, offline experience etc.). Given that there is no guidance on developing one’s own career, it is of high practical relevance to provide digital workers with some sort of career model to judge their own careers and support them in making future career choices.

4. Data and Method
4.1 Sequence Analysis

To analyze our research question, we conduct a sequence and cluster analysis. Scholars have recognized that there is a need to examine complete work histories (Abbott & Hrycak, 1990; Rosenfeld, 1992) because focusing solely on job changes might result in losing sight of the complete career line. However, it is assumed that individuals continually plan and structure their work histories (Abbott & Hrycak, 1990) and as evident the aforementioned study by Leung (2014), the order of complete work histories matter in OLMs.

To analyze sequences, Abbott and Hrycak (1990) have establish the use of optimal matching (OM) to find "typical" career lines, i.e. patterns in sequences of positions. OM is widely known in natural science to measure sequence resemblance, e.g. in examining DNA sequences. The OM algorithm calculates distances between all pairs of sequences in terms of insertion, deletions, and substitutions required to transform one sequence into another (Abbott & Hrycak, 1990). Then, the algorithm measures the minimum number of transformations required (“costs”) to turn one sequence into another. After calculating distances between sequences based on transformation costs, a clustering algorithm is applied to the metric distance matrix. In OM, sequences with low metric distances between each other are interpreted as having similar underlying patterns.
Prior work using OM algorithms has analyzed, for example, the career sequences of musicians (Abbott and Hrycak, 1990), women in finance occupations (Blair-Loy, 1999), and IT professionals (Joseph et al., 2012). We add to this stream by applying sequence analysis to a novel context characterized by several layers of heterogeneity and potential career paths. Thus, sequence analysis represents a very suitable method to identify patterns in OLMs and help understanding whether and how digital workers self-assemble their careers in this context.

4.2 Data

We use an extensive and novel transaction-level data from the platform Upwork to test our hypotheses. Upwork represents an ideal context for our analysis for three reasons. First, Upwork explicitly encourages high-value ongoing work. Studies focusing on high-skill workers in external labor markets is scarce. Further, knowledge worker work on heterogeneous tasks so that some sort of progress is more likely (e.g. through working on more complex tasks over time) and job shifts occur frequently – compared to low-skill, standardized gig work such as driving services, for which career development is difficult. Second, it is a global platform so that we can analyze geographical differences in how workers progress through their careers. Third, Upwork is the largest freelancing website worldwide. This increases the likelihood that freelancers are less likely to multi-home but rather focus exclusively on Upwork given the supply of work.

The original dataset was gathered in 2017 using a Python script and includes data on 255,393 freelancers with a minimum of one job. The sample was then restricted to 71,336 freelancers with a minimum of 10 jobs to ensure a certain career trajectory, resulting in 2,662,983 observations (transactions). The starting dates of the projects in the sample are between March 2006 and October 2017.

We are currently at the stage of coding the data and defining substitution costs, which is a nontrivial task and is crucial for the quality and clarity of produced results (cf. Abbott, 1990).
References


