Mapping Career Paths in the U.S. Labor Market

Marc Scott[†], Matthew Zeidenberg[‡], Annette Bernhardt[*], Laura Dresser[‡]

ABSTRACT:

Using the National Longitudinal Survey of Youth (NLSY), we analyze the mobility paths of young workers' careers between 1979 and 2000. We identify three "mobility groups" in the U.S. labor market: (a) workers who are stuck in low-wage jobs over the long run; (b) workers who start out in low-wage jobs but then managed to escape them; and (c) workers who manage to avoid low-wage jobs altogether. Our focus is on uncovering the structure of low-wage careers, both in terms of what they have in common with, and what distinguishes them from, their more mobile counterparts. Using a novel method of matching and clustering, we are able to construct a meaningful typology of career trajectories based on the sequencing of industries, occupations, and movements in and out of labor market. This typology suggests that the bulk of low-wage careers are relatively stable and show strong industry-occupation patterning, and that job stability plays an important but highly contextualized role with respect to mobility. The typology also allows us to evaluate several quasi-experiments, isolating features that distinguish more mobile and less mobile careers with similar industry-occupation profiles.

6/17/2005

EARLY DRAFT, PLEASE DO NOT CITE OR DISTRIBUTE

This research was supported by a grant from the Annie E. Casey Foundation. The authors wish to thank Heather Boushey, Barbara Gault, Steve Herzenberg, Chip Hunter, Dave Marcotte, Pablo Mitnik, and Julie Strawn for their comments on previous versions of this report. The authors also wish to thank Joel Rogers and David Giloth for their guidance, support and enthusiasm. Research assistance provided by Rebecca Glauber and Erez Lenchner is also appreciated.

[†] Corresponding author. Department of Humanities and Social Sciences, NYU Steinhart School, 246 Greene Street Suite 318E, New York, NY 10003. Email: marc.scott@nyu.edu.

[‡] Center on Wisconsin Strategy, University of Wisconsin, 1180 Observatory Dr. Room 7114A, Madison WI 53706.

[*] Brennan Center for Justice at NYU School of Law, 161 Sixth Avenue, 12th Floor, New York, NY 10013.

1. Introduction

This study focuses on the careers of low-wage workers in the U.S. labor market. It is motivated by both the sheer size of the low-wage workforce – numbering between 30 and 40 million depending on the definition used – and by the mounting evidence that mobility out of low-wage jobs is increasingly difficult. For example, Bernhardt et al. (2001) find that the percent of workers earning low wages over their lifetimes doubled from 15% to 30% among young white males, comparing the 1970s to the late 1980s and mid-1990s. Studies looking at more recent data find similar rates of immobility for low-wage workers (e.g. Boushey 2005; Andersson et al., 2004). The negative consequences of this trend ripple throughout economic, social, and political life (Neckerman 2004); to form relevant social policy, we need to understand the forces shaping the labor market in which low-wage workers compete, and their implications for economic mobility over the life course.

A common intuition is that low-wage careers have very little structure. According to this view, there is a lot of job hopping across many different occupations and industries, one job looks much like the next, and the entire process is pretty much chaotic. But is this really the case? As we will see, there are in fact systematic paths that constitute low-wage careers. Some are bound by industry and occupation, others are not. Sometimes job instability is detrimental, and other times it is too much stability that stands in the way of progress. Sometimes valuable skills and experience are gained on the job, sometimes not. But the upshot is that low-wage careers are not black boxes; rather, they have structure. Unveiling that structure is important both to academic research and to interventions in the labor market, such as career ladder initiatives.

Recent longitudinal analyses have found that the career mobility of workers depends highly on the sectors in which they work. For example, Andersson et al. (2004) find that earnings growth tends to occur in the service sector for women, compared to manufacturing and wholesale trade for men. Strawn and Martinson (2000) find that former welfare recipients who started working in clerical positions earned a fifth more on average, five years later, than their counterparts who started working in retail sales jobs. At the very low end of the labor market, Boushey (2005) finds that industry and union status have significant effects on the probability of moving up. In these examples, the workers involved had roughly similar education and skill levels (or were made equivalent by regression), while it was the industries or occupations that varied.

Our work complements and expands on these studies. Using the National Longitudinal Survey of Youth (NLSY), we compare and contrast the mobility paths of young workers' careers between 1979 and 2000. We use poverty-based thresholds to define three "mobility groups" in the U.S. labor market: (a) workers who are stuck in low-wage jobs over the long run; (b) workers who start out in low-wage jobs but then manage to escape them; and (c) workers who manage to avoid low-wage jobs altogether.

Our focus is on uncovering the industry-occupation structure of low-wage careers, both in terms of what they have in common with, and what distinguishes them from, their more mobile counterparts. Using a novel method of matching and clustering, we construct a meaningful typology of career trajectories – defined in terms of the progression of industries and occupations that workers move through over time, as well as their labor force participation. Analysis of these career trajectories reveals significant differences but also surprising similarities across the three mobility groups. For example, the bulk of low-wage careers are relatively stable and coherent, and show strong industry-occupation patterning. Our typology also allows us to evaluate several quasi-experiments, isolating features that distinguish more mobile and less mobile careers with similar industry-occupation profiles.

This paper is organized as follows. In section 2, we discuss the data and our overall methodology. In section 3, we identify the three mobility groups, estimate their relative size, as well as summarize key defining characteristics. We examine these characteristics in aggregate and then by gender. In section 4, we describe a clustering technique that identifies a set of typical career paths for each mobility group. We then highlight some findings based on this approach. Section 5 presents the matching technique that establishes pairs and triplets along mobility and industry/occupation lines and broadly illustrates their distinguishing features. Section 6 presents several case studies, in which we explore the differences between these matched pairs in greater detail. We draw some preliminary conclusions and suggest avenues for future research in section 7.

2. Data and methods

Our data source is the National Longitudinal Survey of Youth (NLSY). The sample is representative of non-institutionalized men & women in the U.S. aged 14-21 in 1979. This cohort was interviewed every year from 1979-1994, and then bi-annually until 2000, when the group was aged between 35-42. Black, Hispanic and poor whites were oversampled in what are known as "supplemental samples." The poor white supplemental sample was dropped from this analysis because it was discontinued after 1990, truncating the career sequence prematurely. Poor whites are represented in the common, retained sample, and our weights have been adjusted to accurately reflect their proportionate contribution. A supplemental military sample was dropped from our analysis sample as well. The original sample size, including all supplemental samples is 12,686. This drops to 9,763 after the two supplemental samples are dropped. After careful evaluation of patterns of missed interviews, we decided to drop individuals who show a gap of more than four years between any two surveys.[1]

When we refer to "the career," we mean a sequence of 2-digit industry and occupation codes associated with quarterly jobs spanning ages 20-36. We construct 25 unique

[1] The reasoning behind this is based in part on our unit of analysis, which is the sequence of industries and occupations held by each worker. The NLSY carefully "reconstructs" the work history for any subjects who miss interviews, but the reconstruction is limited to the five most recent employers. Five employers are sufficient when the time between current and last interval is short. However, across wide gaps, there is a concern that the occupation reported, which will be the most recent for a given employer, will exclude significant changes in occupation. An example of this is a promotion to management; across wide gaps, the worker will appear to have always been a manager, while in fact this is far from the case. Thus, we adopt a moderately strict criterion for inclusion in the study.

industry and 20 occupation codes that aggregate the 3-digit 1970 Census codes into reasonably homogeneous groupings (1970 rather than 1980 Census codes were chosen because these are the only codes collected consistently over the survey span). For example, all skilled manufacturing industries collapse into one code. In the 16-year age span studied, approximately 500 unique industry and occupation pairings (IxOs) occur.

In the reduced sample, missingness on the key variables of industry and occupations for jobs recorded in the work history is minimal at about 3%. About 100 individual cases have missing industry or occupation information for more than half of the work history, so these cases are dropped. The sample size becomes 7,816, or 80% of the maximum possible. We reweight the sample so that it is consistent with the demographics of the original baseline sample.

The career sequences analyzed consist of industry-occupation pairs for the 64 quarters spanning ages 20-36. Unemployment, enrollment, or time spent out of the labor force are coded into the sequence as well. Each career sequence is summarized in ways that reflect the dynamics of the industry-occupation trajectory over time. For example, we calculate the number of distinct industry pairs witnessed in the 64 quarters. Note that some important information is cumulative or predates age 20. Weeks worked from age 16 onward are accumulated into an experience measure.[2] Education prior to 1979 is included, as is prior presence of children in the household. Permanent wage growth (see below) is based on a slightly different period, age 24-38. The lower age reflects a point in the life course at which most individuals have entered the labor force, and the latter is a point at which most family formation, if it is to occur, has begun.

<u>Mobility group definitions:</u>

For each respondent, longitudinal wage profiles were constructed using inflation-adjusted, logged hourly wages associated with the "CPS" job (the current or most recent job) at the time of each annual interview. Taken as a whole these wages form a profile of growth or stagnation over time. The inflation adjustment is made using the Consumer Price Index research series (CPI-U-RS).

The wage profiles were cleaned of short-term wage fluctuations by substituting predictions from a longitudinal mixed-effects model for the original observations. The new profile can be understood as the permanent wage level over a broad time span. Gottschalk and Moffitt (1994) discuss the theory behind the permanent and transient wage decomposition and provide a methodology for their identification. We use a slightly different methodology described in Bernhardt et al. (2001) and give further details in Appendix A.

We classify each permanent wage as either low or not low using a poverty line threshold. Wages below 1.25 times the poverty line for a family of four (converted from annual income to an hourly wage) in a given year are considered low. In 2002, this is about

[2] Prior experience for older members of the cohort is imputed using an education- and demographic-based model

$10/hour. An additional twenty-five cases were dropped from this analysis due to inconsistencies in the wage profiles, including severely outlying wages.

We divide workers into three mobility groups based on the classification of permanent wage at age 24 and 38. Each individual was classified as belonging to one of the following three mobility groups:

- *"Stuck"*: Wages are low at age 24 and are still low at age 38.
- *"Mobile"*: Wages are low at age 24 but are no longer low at age 38.
- *"Never Low"*: Wages are never low, at either age.

Seventy-nine workers were classified as downwardly mobile, but this category was dropped due to small sample size. The final sample size is 7,712.

3. Aggregate findings

Table 2 summarizes career and worker characteristics for three mobility groups in the U.S. labor market. The rightmost column aggregates each measure across the entire sample and gives some initial guideposts to our data. For example, average permanent wages grew from $9.72 at age 24 to $16.32 by age 38 (in 2002 dollars). Respondents worked for an average of 6.6 employers in the 64 quarters examined, and 11.2% of that employment is in the public sector. African Americans comprise 13.9% of our sample, and 40% of the sample has completed no more than high school by age 36.

More interesting are the breakdowns by mobility group. Under our poverty-based definition, fully 28% of the sample is permanently stuck in low-wage jobs over the career. Another 33% begins working in low-wage jobs, but then escapes them by mid-career. And 39% manages to avoid low-wage jobs altogether, throughout the career. These are striking numbers, showing a significant amount of immobility out of low-wage jobs.

At this point it is worth reiterating one of the basic facts that governs careers in the U.S. labor market: the bulk of life-time wage growth and job changing occurs in the first decade and a half in the labor market (Topel and Ward, 1992). By the mid- to late-30s, workers' career trajectories are largely set. Thus, while we are only measuring wage growth between the ages of 20 and 36, in fact we are capturing much of the mobility that the young adults in our sample will achieve.

Not surprisingly, many of the demographic and education measures that follow differ across the three mobility groups, in ways that we would expect. Stuck workers are more often female, less-educated, African American, and living in the South. Never Low workers are more often male, white, college-educated and living in urban areas outside the South.

The job characteristics are also not surprising in their distribution across the mobility groups. Unionization rates are lowest for Stuck workers, as is the related measure of public sector employment. Similarly for employer-provided training (a critical variable and one we hope to explore in much more depth, since it has shown to be nearly as strong a predictor of wages as education).

The interesting comparisons come when we shift to measures of career structure. For example, time spent out of the labor force and/or unemployed is a very strong marker of low-wage careers – not surprising given similar findings in other studies and clearly an important characteristic that requires analysis. However, as we will see below, this aggregate number hides significant variation in low-wage career trajectories. Many of these trajectories show significant labor force attachment, and the majority does not show the type of chronic unemployment or time out of the labor force that we usually associate with low-wage work. Similar findings obtain for the incidence of part-time work, year-round work, and cumulative work experience. On average there clearly is a strong correlation with mobility group, but once we look in detail at typical career trajectories within a given mobility group, there is significant variation.

Even more interesting are the findings for numbers of employers, industries, and occupations. These measures are actually strikingly similar across the three mobility groups. In particular, they are nearly identical between Stuck and Mobile workers, suggesting that job hopping and churning is less of a culprit in trapping workers in low-wage jobs than commonly thought.

Again, in part, these aggregate measures obscure variation in how the structure of careers develops over time. But as we will begin to document below, it also the case that the three mobility groups are often differentiated more by the content of the careers – the actual industries and occupations navigated – than by their attachment to the workplace or the labor force.

Gender-based summaries:

Table 3 breaks down the above measures by gender, as backdrop for the very strong gender differences in career trajectories that we will see below. While time spent out of the labor force tends, on aggregate, to characterize the careers of women, this is by no means uniform. (In section 4, we will demonstrate that many female careers contain highly levels of labor force attachment.) Unemployment rates are higher for Stuck men than for Stuck women, suggesting different challenges for each group in this part of the labor market. Stuck male workers, but not female workers, are more often African American than would be expected.

Other differences include the rate of part-time work, which is about twice as high for women as for men (for all mobility groups). Unionization appears to play a bigger role in delivering good careers for men than for women (i.e. see the Never Low category). Finally, women consistently have higher rates of educational attainment, in all three mobility groups.

4. Clustering career sequences

The above comparisons of the three mobility groups are useful, but raise many more questions than they answer. The root of the problem is that there is simply too much variation in how the structure of a career develops over time – in the *temporal sequencing of jobs and labor market participation as a worker moves through the labor market*.

Our intuition is that we will learn something by treating this sequence of jobs as the unit of analysis. Using a novel method of matching and clustering, we construct a typology of career trajectories – defined in terms of the progression of industries and occupations that workers move through over time, as well as their labor force participation. Analysis of these career trajectories reveals very strong industry-occupation patterning, for all three mobility groups including those workers stuck in low-wage jobs over the long run.

Methodology

As a reminder, the career sequences consist of industry-occupation pairs (such as retail/sales clerk) for the 64 quarters spanning ages 20-36. Quarters spent unemployed, enrolled, or out of the labor force are given unique codes, and form an important part of the career structure. We develop a career typology for sequences identified as Stuck, Mobile or Never Low, by placing those with similar IxO components and structure into the same cluster, and we do this separately for each mobility group. This clustering problem is challenging, as there is no natural metric for comparing two sequences. How do you compare three years as a secretary in a small detective agency (for one worker) to ten years as an auto mechanic at a dealership (for another)? We must consider how important the timing of the job is to the matching process. At one extreme, one could align the two sequences quarter-by-quarter and assign a sequence of 1s and 0s, where a 1 is given only when the IxO pair across the two workers matches exactly.[3] The average of these assigned values across the 64 quarters reflects how far one career is from another, using this very conservative matching criterion. At the other extreme, one could ignore the ordering of these sequences and construct a dissimilarity index based on the overlap in the *distribution* of industry-occupation pairs across two workers (see Massey and Denton, 1988, for a discussion of dissimilarity indices). Under the latter distance metric, a career that is half in retail sales and half doing accounting services is the same as any other career split between these two IxO, *no matter what the ordering*.

The challenges in the clustering problem can be divided into two subproblems. First, we need a more nuanced metric for comparing the IxOs across two sequences. The IxOs are the fundamental building blocks of sequences, and they provide us with an alphabet in the same way that the letters A-Z build words or sentences. We want to compare sentences (our sequences), but to do so, we need to know how "far" an 'A' is from a 'B.' Thus we are forced to decide how near or far each IxO pair is from another. For example, we need to know whether a "restaurant/waiter" job is near a "department

[3] A more nuanced approach would assign the value 0.5 to a half-match when only the industry or occupation (but not both) matches.

store/sales clerk" job. Second, we have to decide how to align two careers so that the token-by-token pairings may be compared and an overall distance computed. For example, what should we do if two careers in healthcare differ in a small way, such as one being preceded by a year of restaurant work, while the other entered the healthcare industry immediately?

To answer these questions, we establish the following set of guidelines for a meaningful comparison of two sequences. Our choices for metric and alignment should ensure that:

1. Two sequences that contain the same modal (most frequent) token, when it truly dominates a substantial portion of each sequence, are deemed similar.
2. Two sequences that are comprised of the same set of tokens occurring at the same rates are similar. That is, careers that are a mix of retail/sales and FIRE/clerical in about the same proportions are similar. Careers are more similar if timing of the different tokens matches reasonably closely.
3. The point at which a particular token occurs in a sequence is treated as if it were somewhat imprecise (i.e. we suspect that labor market activity at age 23 is comparable to age 24, so sequences are still considered similar, even when they are not *perfectly* coincident).
4. A few unexpected tokens in a sequence may be ignored. Infrequently occurring tokens may be taken as "noise" (and is thus ignorable) for comparison purposes.

These guidelines are effectively met by using an optimal matching algorithm (OMA; see Abbott, 1995). OMA effectively cleans a sequence of non-representative jobs before making comparisons and it allows for minor shifts and gaps in careers, all of which are suggested by (3) and (4) above. It also may be calibrated to use a distance metric that reflects principles (1) and (2) above. We use a distance metric based on the conditional probability (within a sequence) of different IxO pairs. Intuitively, if a long stretch of durable manufacturing/operative work is present in nearly every career that contains some manufacturing/manager jobs, then the probability of the latter occurring, given the presence of the former is reasonably large. We set the distance between two IxOs to reflect this probability. The choice of metric and some further details on our implementation of OMA are discussed in Appendix B.

OMA generates a set of distances between every pair of career sequences. These are then input to a clustering algorithm that places 'nearby' careers together in the same cluster. Again, we do this clustering separately for each mobility group. The sequences in each cluster typically share a core set of industries and occupations; they also share similar structure, such as significant early or late departures from the labor force.[4] We view the ensemble of career clusters as a typology, or a way to label similar careers based on their component industry-occupation (and time spent out of the labor force) trajectories. We partition the careers into clusters using the Partition Around Medoids (PAM) algorithm (Kaufman and Rousseeuw, 1990). Within each mobility group, PAM classifies each career sequence into one of sixteen categories. The choice of 16 career clusters within

[4] As a reminder, tokens for education, unemployment and OLF periods are part of the career sequence.

mobility group was based on a global goodness of fit statistic known as the silhouette width.

Results

Table 4 summarizes the career clusters that we have identified for each of the three mobility groups. This is a bird's eye view, and is meant simply to give a flavor for the types of clusters identified, and the percent of workers in each. In particular, we should emphasize that the cluster labels are simply short-hand, and should not obscure the considerable nuance and diversity of career paths within clusters.

We illustrate this point by displaying several actual clusters we have been analyzing. Exhibit A provides a visual depiction of our durable manufacturing cluster in the Stuck and Mobile groups. On each page, a random sample of individual career paths from that cluster is shown. Even without the legend (attached at the end of the exhibit), it is clear that while trajectories in a given cluster share some core characteristics, there is in fact a good amount of variation around that core, and often that variation has a structure of its own.

In order to manage the large amount of information inherent in these clusters, we recompute the summary measures that capture key features of the career sequences on a cluster by cluster basis. These features include the patterning of labor force participation, the changing mix of industries and occupations, and other indicators of career structure. Each statistic is calculated across the entire career, and then also for two sub-periods in the career to give us a sense of the timing and magnitude of these key features. Integrating both the raw data (Exhibit A plots) and the summary measures, we construct "summary profiles" for each of the 48 clusters.[5] In Exhibit B, we show eight very different examples to give the reader a flavor of how rich these clusters are – and also how very structured they are.[6] Some technical details relating to the symbols used in these plots is given at the end of Appendix B after presenting how they were derived.

These profiles represent the variety of career types we witness. Moving through Exhibit B sequentially, Stuck cluster #2 (this is an internal index that can be linked to Exhibit A plots), secretaries and bank tellers, reflect highly attached, female dominated, Stuck careers in the FIRE (Finance Insurance and Real Estate) industry. One feature of this career type is extremely little job changing early on.

Clearly, even the summary versions of each cluster represent a substantial amount of information, and we are still in the process of digesting the many lessons and insights they yield. At the end of this report we describe our plans for statistically assessing the

[5] Our strategy in summarizing each cluster is to first identify the dominant career path that lies at the core of the cluster, then to identify variant paths, and finally, to search for any evidence of occupational mobility in the cluster. It is important to note that although we specifically seek to identify coherent structure in these career sequences, we cannot impose order on chaos. If a significantly large cluster in the Stuck group emerges that is highly attached to the labor market, it must be a feature of many of those careers.
[6] Four of the eight (durable manufacturing and healthcare) were chosen because they will be discussed in greater detail in section 6.

factors associated with cluster membership. But in the interim, it is worth reviewing some of the broad trends that our ongoing analysis has yielded:

- Looking across the mobility groups in Table 4, a surprising number of industry-occupation paths show up in more than one mobility group. For example, there is a distinct retail cluster in each mobility group, and these are actually quite similar on many dimensions (except, of course, wages). Other examples are education, secretarial work, restaurant work, and manufacturing. Analyzing these shared clusters side-by-side is a powerful way to identify, without statistical models, some of the factors that determine mobility group membership. In effect, these are quasi-experiments, in which many aspects of the career are controlled via the matching. We expand on this idea in Section 5 via two case studies of such cross-mobility group comparisons.
- Many Stuck workers show a substantial amount of coherence in their career paths – staying in one industry and one occupation for sustained periods, or making logical progressions from one type of job to the next. Good examples are "stable retail career" (detailed profile not given) and "secretaries and bank tellers". Many also show job stability and strong labor force attachment (though often tempered by part-time work). Put another way, the number of Stuck clusters that exhibit chaos is quite small – while a sizable minority do contain large stretches out of the labor force, what remains in those careers is overall quite consistent.
- Regarding time spent out of the labor force, the cluster reveals a bifurcation, somewhat strongly upon gender lines, of Stuck careers that are weakly attached and those that are strongly attached. There is less "middle ground" here than one might suspect. In fact, when our sixteen Stuck career types are divided once again by gender, the resulting 31 clusters have a 50% larger variance than the comparable careers in the Mobile group.[7]
- There is a rough correlation, on aggregate, between education and the three mobility groups (see Table 2). Yet educational distributions within the individual clusters are surprisingly quite broad. The result is that when comparing individual clusters to each other (especially Stuck ones with Mobile ones), educational differences are often at the margins only. The same holds true for amount of training provided by the employer.[8]
- There is a pronounced union effect, which came as something of a surprise. The effect plays out in a number of ways – both directly, by boosting the wages in a given industry-occupation cluster so that the cluster lands in the Mobile or Never Low group, and indirectly, by increasing the amount of training and promotions that

[7] The idea here is that clusters built to reflect coherent industry-occupation structure are more diverse in the Stuck group, with a solid majority whose sequences look very much like any other group's juxtaposed with a set of more extreme career types, such as those that are dominated by time spent out of the labor force. There are 31 rather than 32 unique clusters because there are no Stuck female "auto mechanics and repairmen".

[8] In fact, employer-provided training, while on average much less common for Stuck workers, matches Mobile workers' training levels about half of the time, when evaluated on a cluster-by-cluster basis. The same finding holds for education, only a bit less so.

workers receive. There is also an interesting illustration that unions continue to be able to reduce wage inequality in a few sectors: several Never Low clusters are bifurcated between high school graduates (typically unionized), and college graduates who are capturing the returns on their credentials.

- Another pattern that is rich for analysis: some industry-occupation combinations form their own career clusters, but then also are feeder jobs for other careers – retail sales jobs are the classic case here. Understanding why retail jobs sometimes lead to solid careers and other times do not is a key question. Judging by our analysis so far, it is not primarily a matter of education; industry segment and employer investments in training seem to be stronger candidates.
- Sometimes, the difference between Stuck and Mobile clusters is precisely that: the latter show mobility from entry-level to managerial or technical occupations. But we also found some evidence of occupational mobility *within* Stuck clusters – for example, movement from retail clerks, secretaries, or bank tellers to manager.
- Gender segregation at the cluster level is highly pronounced, both within and between mobility groups. Racial segregation is also evident, though not quite as strongly. In a preliminary analysis, we find that Stuck worker career types are more segregated than the remaining two mobility groups. About half of the gender-specific Stuck clusters are about a third or more people of color. That figure drops precipitously in the more mobile groups. In other words, there are no Mobile or Never Low career types that can be characterized as dominated by people of color.[9]
- One interesting lesson from comparing predominantly female clusters across mobility groups: part-time/part-year work is not necessarily a negative. Significant numbers of women manage to develop well-paid careers without full-time, year-round work.
- We do have several sizeable clusters that consist largely of time out of the labor force. But more can be discovered beneath that label. In particular, the clusters labeled "infrequent service jobs" (in the Stuck and Mobile groups) show some patterning in the mix of industries and occupation and how the exits from employment are patterned.

These are clearly only initial observations, and they are largely generic. In what follows, we use two case studies to dig deeper into an industry-occupation specific analysis.

5. Establishing matched pairs

As previously noted, it is interesting that in Table 4 so many similar sounding industry-occupation paths show up in multiple mobility groups. Take, for example, the retail sector with substantial Stuck (stable retail careers), Mobile (retail hoppers), and Never Low (retail managers and sales) clusters. By looking more closely at jobs and their

[9] Postal and other government jobs, as well as unionized utilities have the largest share of non-white workers in the Never Low group. Truck and bus drivers and unionized non-durable workers for Mobile workers of color stand out as well.

attributes in these seemingly similar clusters, we hope to learn more about the ways that specific job attributes contribute to mobility patterns.

Review of Table 4 provides some cursory evidence of the ways that clusters with different mobility can share very similar industry/occupation patterns. To formalize such observations, we developed a method to identify matched pairs where industry and occupation clusters look similar, but mobility is quite different.[10]

Table 5 summarizes the results of this analysis, showing relationships between clusters that can be linked across mobility groups on each row. Clearly many such contrasts exist, and understanding the unique attributes and quality of jobs which distinguish these pairs will help illuminate key questions of mobility. We make such case study comparisons in the next section.

Before moving to the case studies, however, it is worth noting that along with the set of careers that are similar across mobility lines shown in Table 5, there are several that are quite isolated. For example, Stuck hairdressers and childcare workers do not seem to share industry-occupation structure with any more mobile cluster. Moreover, this cluster shares little with the other clusters in the Stuck group, making even a prolonged path out of low-wage work (through a Stuck intermediary conduit) unlikely. The Never Low extreme, Computer Programmers and Engineers, is similarly isolated. Even though the bulk of the work is in Durable Manufacturing, there is no apparent career line linking these to any other careers (see Spenner et al., 1982 for a discussion of career lines).

6. Analyzing contrasting mobility in similar industry-occupation clusters

Why do such similar looking careers produce such disparate trajectories? In order to contrast and characterize these matching industry/occupation clusters, we first closely review the summary profiles described previously. These establish both worker and firm based differences, such as level of education, training, or unionization. We pay close attention to the career sequence summaries, such as increased levels of unemployment or increases job and/or industry-occupation switching (these are not necessarily the same) over time. These may suggest deunionization or other restructuring of a subsector of the primary industry, or a local economic downturn.

When matched pairs are very similar in terms of career sequencing, more in-depth analysis is needed. For example, clerical and administrative work in the FIRE industries is usually highly attached, career-wise, no matter what mobility group is considered. Something other than unemployment and job-switching is at play in this industry. For

[10] Formally, we construct a dissimilarity index that matches jobs at the 2-digit IxO level across mobility groups and identify pairings or triples representing substantially similar careers. Our requirement is that these matched pairs have at least one half of their sequence components matching at the 2-digit IxO level. Triplets were established when at least two of the three possible cross-mobility comparisons met our threshold for similarity. See Appendix C for more details on the matching process and criteria.

healthcare, explanations based on education might emerge, but the timing of the education, which has clear links to socioeconomic status, seems to be important as well. In manufacturing, we will see that the product lines make a big difference: building cars pays much better than building furniture, although the entry-level skill set for assemblers may be comparable. In construction, working in the unionized part of the industry has the most substantial pay-off.

We have yet to fully address the multitude of gender, race/ethnicity and perhaps regional explanations for differing outcomes, and the matched pairs analysis is a good place to do this work, as so many other aspects of the career are controlled with this design. However, "discovering" that it is "bad" to be female if you work in non-durable manufacturing seems minimally useful at best. Of course, we want to know who is experiencing what in the labor market, but the context –jobs—that surround this experience are much more important to us.

While there are many interesting case studies to explore, we concentrate here on an example from healthcare and from manufacturing. Healthcare is a growing portion of the overall economy, with a broad range of subspecialty occupations ranging from patient work including attendants and nurses, a large set of clerical and administrative jobs, and more technical work in the lab. Education, particularly at the sub-baccalaureate level, plays an important role in defining occupations and promoting advancement. Manufacturing has been in decline in terms of workforce share, but it still represents a fundamental career type, comprising nearly 20% of the jobs of the workers we study (it is a smaller portion of the overall economy). For the manufacturing case study presented, we concentrate on durable manufacturing (a canonical point of reference).

Determinants of Advancement in Health Care: The Roles of Education, Unionization, and Industrial Subsector

In what follows, we compare two clusters, one of Stuck workers, one of Mobile workers, both of whom are found mainly in health care. Some of the data characterizing these two groups are given in Table 6. We now compare the characteristics across mobility groups.

Both clusters exhibit delayed entry into the labor market. The Mobile group does not move definitively into nursing until their late 20s. Some of the Mobile go from being nursing aides to nurses or technicians.

The Mobile group has somewhat more weeks of labor force experience by age 36 (766) than do the Mobile (679), so the increased amount of time spent outside the labor force may be a factor contributing to the failure of the Stuck to advance. However, there is very little difference in the share of time spent working part-time (22% for the Stuck, 24% for the Mobile, across all ages) The Stuck have more employers, 7.5 to the Mobile's 6.6 (again, across all ages), indicating that they are churning a bit more, moving from one low-wage job to another in a variety of industries besides health care, while the Mobile often stay in a higher-wage job if possible once they get it.

There is a difference between the groups with respect to unionization. The unionization of the Stuck group declines from 17% in the early 20s to 14% by the mid-30s. Over the same period, the Mobile workers' unionization was rising, from 15% to 21%. Thus the Stuck workers are losing or are unable to get union jobs, and the Mobile workers are holding on to them and/or obtaining new ones. Of course, this usually leads to better pay.

The education gap between the groups is large. Among the Mobile, 53% have at least some college by age 25, as opposed to 24% among the Mobile, a 29 percentage point gap. This narrows slightly to a still substantial 27% by age 36; 63% to 37%. Thus education may account for much of the advancement story. However, even the Stuck overwhelmingly have at least a high school education (88%, as opposed to 96% of the Mobile).

Healthcare is one of a few sectors that we found an important role for sub-baccalaureate education. Of the Mobile workers, 20% obtained an en route (or terminal) Associate's degree (Licensed Practical Nurses are likely to pursue this), while only 7% of Stuck workers do. Baccalaureate attainment is important, but seems to play less of a role in this segment of Healthcare than it often plays in other fields, for which a bachelor's is minimal requirement.

There is not an enormous difference between the groups in the amount of training that they receive (the Mobile get a bit more in their late 20s and mid-30s) nor is there much difference in where they live. The latter implies that we are not just witnessing depressed wages in rural or southern communities, for example.

Both groups are overwhelmingly female—the Stuck slightly more so (82% versus 75%). Thus it does not appear that gender is a major determinant of advancement. There is little difference in the share of the groups that are black (28% for the Stuck, 22% for the Mobile). Both groups are more black than is the general population. Among the Stuck, 8% are Hispanic, as opposed to 5% of the Mobile.

There is not much difference between the groups in the number of children that they have in their household. The Stuck have on average 1.4 children by age 36, the Mobile, 1.3 (figures not reported in table). The Stuck have a bit more children on average in their early 20s; this may have something to do with the education gap and increases time out of the labor force early for the Stuck.

<u>Industries</u>

There is a diversity of wages paid in different industries within health care. In 2001, the Occupational Employment Statistics program reported a median wage across all workers (not just health care) of $13.01. This wage was $14.50 in doctors' offices, $10.11 in nursing homes, $16.35 in hospitals, and $10.18 in home health.

In what follows, we first match the industrial, occupational, and industrial-occupational distributions for the two mobility groups. For the purposes of example, we will concentrate here on the distribution of industries. We first report those industries that are more common in Stuck workers than in their mobile counterparts. They may not comprise a particularly large share of Stuck work, but they stand out as the more prevalent as compared to matched Mobile workers. We then report a similar list for the Mobile group, which is then followed by the bulk of the jobs that they have in common.

Stuck workers are more likely than Mobile workers to be found in the following (top five) industries:

- Convalescent institutions (nursing homes)
- Eating and drinking places
- Grocery stores
- Welfare services
- Elementary and secondary schools

These are all relatively low-wage industries, except for elementary and secondary schools, and even this latter industry has low-wage jobs in it. Stuck workers are more likely, if they are working in health, to work in nursing homes, which is a low-wage part of health, as we have seen. Thus we see that the particular subsector of health care in which they work can be a major determinant of their fate. They also spend some time working in other common low-wage industries, such as bars, restaurants, and grocery stores. This may contribute to a lack of tenure in health care that bars their advancement.

Mobile workers are more likely than Stuck workers to be found in the following (top five) industries:

- Hospitals
- Offices of dentists
- Offices of physicians
- Health services, not elsewhere classified
- Department and mail order establishments

In contrast to Stuck workers, we see that Mobile workers are more likely to stay within health. The first three of these industries are relatively high paid subsectors of health care; even the fourth one has an OES median wage of $14.11.

The top five industries in terms of the overlap between the two groups are:

- Hospitals
- Convalescent institutions
- Health services, not elsewhere classified
- Offices of physicians
- Eating and drinking places

It is not surprising that hospitals, nursing homes, and doctors' offices appear on this list. We know that these employ large numbers of health care workers. Each of them has large diversity in wages, which can account for why they can employ both Mobile and Stuck workers. For instance, the 10th percentile wage in hospitals is $8.47, the median is $16.35, and the 90th percentile is $30.24. Large hospitals can be like little cities in themselves, with internal labor markets, a wide variety of occupations, and opportunities for advancement. On the other hand, in nursing homes, the bulk of employment is relatively low-paid nursing aides.

Occupations

Stuck workers are more likely than Mobiles to be found in the following (top five) occupations:

- Nursing aides, orderlies, and attendants
- Billing clerks
- Laundry and dry cleaning operatives, not elsewhere classified
- Key punch operators
- Secretaries, not elsewhere classified

These are all low-paid manual or clerical occupations. Examination of the career tracks of Stuck workers shows a good deal of occupation and employer changing. Some nursing aides become technicians, clerks, or practical nurses, but there is not much occupational mobility in general among the Stuck.

Mobile workers are more likely to be found in the following (top five) occupations:

- Registered nurses
- Physicians, medical and osteopathic
- Practical nurses
- Dental assistants
- Health technologists and technicians, not elsewhere classified

These are all high-paid, skilled occupations that require at least some college education, and four-year or postgraduate study. We have seen that the Mobile have substantially more education than the Stuck.

The top occupations in terms of overlap between the two groups are:

- Nursing aides, orderlies, and attendants
- Health aides, exc. nursing
- Clinical laboratory technologists and technicians
- Practical nurses
- Health technologists and technicians, not elsewhere classified

These are all low or middle-level occupations whose wages can vary as a result of other factors, such as union status or region of the country. For instance, nursing aides can have variable salaries depending on the setting in which they work, such as a government-run versus a private for-profit nursing facility.

Determinants of Advancement in Durable Manufacturing: The Roles of Unions and Career Stability

In this section, we compare two clusters, one of Stuck workers, one of Mobile workers, both of which have workers who are most frequently found in durable manufacturing.

In the Mobile cluster, workers either enter durable manufacturing directly, or pass through non-durable manufacturing. Workers in the Stuck cluster tend to enter durable manufacturing directly. Some of the data characterizing these two groups are given in Table 7. We summarize salient differences between the two groups in what follows.

These two groups of workers are almost identical in terms of their attachment to the labor force: by age 36, the Stuck have 791 weeks of work experience, as opposed to the Mobile, who have 788. So attachment cannot explain mobility. Both groups are working a high proportion of the time, even early on. They also have almost identical numbers of employers over the period from ages 20 to 36; 6.4 for the Stuck and 6.7 for the Mobile (aggregate numbers—not in table).

However, Stuck workers actually work more than Mobile workers in the beginning of their careers. But this does not pay off for them in the long run. Later on in their careers, they are relatively less attached (compared to the Mobile) to the work force and are working in less desirable jobs. Thus the ability of the Mobile workers to find a good job and hold on to it later on in their careers appears to be a major contributor to their success.

One striking contrast is that the Mobile workers' unionization goes up steadily over their careers (from 14% in their 20s to 23% in their 30s, as opposed to 20%, and 11% respectively among the Stuck). Thus over the same period, the Mobile were finding union jobs, and the Stuck were losing them. This appears to be a large factor in their mobility.

There is a substantial difference in the percentage receiving training, with the Mobile receiving almost twice as much employer-provided training. Thus some of the Mobile are in jobs in which their employers are investing in them, which may be boosting their wages.

The Stuck are substantially more female (40% v. 22%) and black (17% v. 10%) than the Mobile. The two groups have roughly the same proportions of Hispanics, the Mobile with a bit more (Stuck: 5%, Mobile 7%). With respect to women and blacks, this pattern is consistent with a dual labor market, to some degree.

The Stuck are more likely to be found in the South than the Mobile. About 50% of the Stuck are in the South, as opposed to 29% of the Mobile. The South has traditionally been a lower-wage region; this is in part due to the relative dominance of lower-wage industries and in part due to lower unionization. In addition, the Stuck are slightly less likely to be in urban areas (60% versus 69%).

There is not much difference between the two groups in terms of how many children they have. The Stuck have 1.3 children by age 36; the Mobile, 1.4. There is also little difference in what proportion of them have children by this age; 65% for the Stuck, 68% for the Mobile.

There is not much difference in the education of the two groups, the Mobile being slightly more educated. Over half of the Stuck and the Mobile have high school degrees (62% for both groups), The Stuck do have more slightly more high school dropouts (23% versus 19%), and the Mobile slightly more college attendees, a few of whom even have four-year degrees.

Industries

Stuck workers are more likely than Mobiles (see discussion framing this comparison for Healthcare) to find themselves in these (top five) following industries:

- Sawmills, planing mills, and mill work
- Furniture and fixtures
- Eating and drinking
- Fabricated structural metal products
- Electrical machinery, equipment, and supplies, not elsewhere classified

The median wage for all workers in the 2001 Occupational Employment Statistics (OES) survey was $13.01. The median wage in sawmills and planing mills was $11.23. The median wage in furniture and fixtures was $11.74; in eating and drinking places, $7.01, in fabricated structural metal products, $13.52. For the fifth group there does not appear to be an exact match with an OES (SIC) industry. But we see that the top three all pay below-average wages, and the Stuck workers must be in the lower tiers of the two bottom industries (since all industries show a significant wage spread).

Eating and drinking is an industry that tends to employ workers who have also worked in a wide variety of industries. Since the Stuck are working in eating and drinking more often, they may be having more difficulty gaining seniority in a job that would protect them from a layoff and perhaps allow them to become Mobile (e.g. by moving up to foreman). We have also noted that the Stuck often work in other parts of retail. Perhaps too much of their careers are taken up in these other industries; although they are working, from the point of view of their careers in durable manufacturing, they might as well not be working when they are working in a bar or restaurant or other retail job.

Mobile workers are more likely than Stuck to find themselves in these (top five) industries:

- Machinery, except electrical, not elsewhere classified

- Other primary iron and steel industries
- Special trade contractors
- Construction and material handling machines
- Metalworking machinery

Machinery production (the first, fourth, and fifth industries listed above are forms of this) is a high-value-added industry, and therefore tends to pay high wages. According to the OES, the median wage for industrial and commercial machinery and computer equipment was $16.18; the three machinery industries listed above fall into this category. The closest match to "other primary iron and steel industries" is the SIC code "miscellaneous primary metal products;" the wage in this industry is $14.42. The median wage for special trade contractors is $15.76. Thus we see that all the industries that Mobile workers are more likely to be in are relatively high-wage, or at least above average.

Mobile workers may be taking their skills from manufacturing and using them in skilled construction labor during layoffs. This contrasts sharply with Stuck workers, who have received less training, and often work in bars and restaurants when they are not working in durable manufacturing.

The top five industries in terms of overlap between the two groups are:

- Motor vehicles and motor vehicle equipment
- Electrical machinery, equipment, and supplies, not elsewhere classified
- Furniture and fixtures
- Logging
- Miscellaneous manufacturing industries

Thus it appears likely that the stratification within these five industries is accounting for some of the differences in mobility. In particular in automobile manufacturing, lower wages can be found in some of the parts suppliers and in the South. There is also within-firm stratification. While the median wage in the automobile industry is $21.86, the 10th percentile wage is only $9.91.

Occupations

The Stuck are more likely than the Mobile to find themselves in the following occupations (top five):

- Assemblers
- Punch and stamping press operatives
- Fork lift and tow motor operatives
- Lumbermen, raftsmen, and woodchoppers
- Freight and material handlers

These are occupations that require relatively little training and have no supervisory role. Except for lumbermen, these are occupations that exist in many manufacturing industries.

For instance, our data show us that assemblers and fork lift/tow motor operators exist in furniture and fixtures, which is a relatively low-wage industry in which the Stuck are also more likely than the Mobile to find themselves.

The Mobile are more likely to find themselves in the following occupations (top two):

- Foremen, not elsewhere classified
- Managers and administrators

Movement into such jobs clearly accounts for some of the increased mobility in the Mobile group. The number of foremen among the Mobile doubles over a ten-year period. However, the Stuck group is not completely bereft of these occupations. Of course, there are situations where one can be a foreman or a manager and one still does not make a good wage.

The top five occupations in terms of overlap are:

- Assemblers
- Machine operatives, miscellaneous specified
- Truck drivers
- Welders and flamecutters
- Checkers, examiners, and inspectors; manufacturing

Assemblers have been a dominant group within durable manufacturing, as have been machine operators; these are both large groups and there is variability in their wages. Assemblers in furniture and fixtures are likely to make less than assemblers in motor vehicles. Truck drivers can show substantial variability in wages; contrast a driver for a small local delivery service with a UPS driver organized by the Teamsters. (Even within UPS, there have been large gaps in the wages of part-time and full-time drivers). Welders are common as well, more so among the Mobile, since they are more skilled. Inspectors, etc. can show substantial variability in wages, based on the particular industry in which they work and their union status, among other factors.

Some general comments on the remaining case studies

As we have just learned, unions have been an important part of mobility opportunities for young workers in the U.S. We have yet to examine the Never Low group in much detail, but unions are strong in more than 40% of the careers we identify (unionized durable and non-durable manufacturing, unionized construction, utilities, government, education and healthcare). The role of unions is by no means straightforward in these and their counterparts in the Stuck and Mobile groups. And there are many parallels between manufacturing and retail in terms of within-industry advancement and firm investment in worker training, but unions are nearly absent from retail, so a different set of investments and conditions must be at play. Credentials play an important role in education and healthcare, and historically government jobs developed clear internal career ladders through a system of exams and job "grading."

We have yet to discuss circumstances in which there is a clear delineation between industry stories and occupation stories. We would like to know when changing industries leads to more mobility, versus when changing occupations does. Since healthcare is such a large and somewhat specialized field, we do witness some level of occupational segmentation in health careers. The technician job seems to be the only common thread between Stuck healthcare workers, such as certified nurse's aides, orderlies and attendants, and the more credentialed nurses and dental assistants. Never Low healthcare work consists of nurses who complete their education early on, and doctors, whose path to high-wage earning is reasonably straightforward. Medical receptionists form a clear mid-range segment of the healthcare industry. Occupational change is rare out of these very stable careers, beyond promotion to office manager. We think that agricultural jobs, teaching, construction, and restaurant work are fairly segmented markets, with a different set of barriers to advancement that should be uncovered.

We have not thoroughly developed our understanding of gender, family responsibilities and mobility, and we have opportunity to do so with the matched pairs involving service jobs. Weak attachment to the labor force is sometimes a clear impediment to mobility, yet this is not always the case. In our Never Low early labor market exiters cluster, we have preliminary indirect evidence that women who postpone childrearing and obtain college-level schooling early on are able to move in and out of the labor force without penalty. The role of socioeconomic status (family resources) clearly needs to be understood in this case.

7. Future directions

Clearly this research project is very much in progress. We have developed a novel methodology that, to our knowledge, is the first to treat *careers* as the unit of analysis in examining the determinants of labor market mobility. The quantity and depth of information that this clustering method yields is substantial (and we have only started to harness it).

The next phase of analysis will use statistical models to yield more precise statements about the factors that drive career mobility – with an eye toward the role played by what we might call path dependency in the trajectories of industries and occupations that workers build over time. A variety of statistical methods could be used to analyze the relationship between industry-occupation trajectories, unemployment spells, unionization, employer-provided training, education, race, gender, and other variables. An important part of this analysis will be to understand the structural constraints on upward mobility – the extent to which there are enough slots to absorb workers trying to move up the ladder. Mitnik and Zeidenberg (2004) discuss how the employment structure of various industries—that is, the relative numbers of low- and high-wage jobs within them—place serious constraints on the possibilities for advancement. They use

CPS data to show that the chances of advancement vary substantially across a selected set of industries.

More generally, we think of the process whereby workers are trapped in low wages as a probability tree, where different combinations of jobs and choices gradually push individuals onto different wage trajectories over time. The "fanning out" of those trajectories has been well documented. But the analogous process of career divergence at the industry and occupation level, and the actual *mechanisms* that drive this process, still require deeper analysis. There is currently great interest in the issue of career mobility, particularly for low-wage workers, and the data, measures, and techniques developed in this project are well suited to the task.

Appendix A. Permanent wage definition

[Appendix incomplete]

Rough outline:

1. Annually reported “CPS job” wages were adjusted for inflation using the CPI-UR-S series.
2. A linear mixed effects model using age as the underlying timeline, a quintic mean structure and quadratic random effects proved the best fitting model for the variation observed. Year-specific variances allowed for secular trends in wage instability. Bayesian Information Criterion (BIC) was used to select the model.
3. Predictions for every age made using the best linear unbiased predictor (BLUP) of the outcome given the model and an individual’s data.
4. Wage profiles with extreme multivariate outliers were removed from the analysis (25 individuals).

Appendix B. Clustering the career sequences

Optimal matching techniques (OMA) were introduced by Sankoff and Kruskal (1983), applied extensively in the social sciences by Abbott (1995), and are now commonly applied to Biological Sequence Analysis (Durbin, et al., 1998). OMA is an algorithm that computes a distance between any pair of sequences, finding the minimal set of primitive operations (insertion, deletion, and substitution) that transforms the source sequence into the target. For example, sequence 'AABBC' can be transformed into 'AAABB' by either inserting an 'A' before the 'B' and then deleting the 'C'; or by substituting 'A' for the third 'B' and then 'B' for the final 'C' (other possibilities exist). OMA finds the set of operations with the minimal cost, where each primitive is assigned a specific cost. Typically, insert and delete are given fixed costs, while substitution depends on what is being substituted.[11]

Token distances

OMA is a useful method for clustering, in that sequence pairs are assigned distances in a deterministic and optimal fashion. We can then use these in a clustering algorithm that takes dissimilarities as input (see Appendix C for details). However, significant user input is needed to run an OMA successfully. In particular, the cost matrix associated with substitution plays a major role in determining the dissimilarities and thus clusters. At one extreme, the analyst could make the substitution cost zero for tokens A and B, effectively merging them together. How surprised would we be to discover that sequences 'ABBB' and 'BAAA' were zero units apart? At the other extreme, all substitutions could be assigned cost equal to the sum of the insertion and deletion costs, ascribing no advantage to substitutions. The latter is a default set of costs for some OMA programs. Vingron and Waterman (1994) discuss these and related issues in the context of biological sequences.

Thus, implicit in the OMA approach is the relationship between the tokens, or components of the sequence. Ideally, we would like these relationships to emerge from some analysis of the sequences themselves, rather than being imposed arbitrarily. But one of the big challenges with categorical data is that there is no notion of distance: how far should a manufacturing skilled laborer be from a professional healthcare worker? If the tokens were to "reside" on some physical space, such as the two-dimensional plane, in which distance between tokens reflects their similarity, these distances could be used to assign substitution costs. This latent distance approach has been used in Social Network Analysis (see Hopf, Raftery, and Handcock, 2002).

We take a more direct approach and employ a substitution matrix for OMA that is suggested by the literature on biological sequence analysis. If the joint probability, $P(ab) \equiv$ the probability of witnessing token 'a' and 'b' together in the same career sequence, is "large" in some sense, then these tokens are similar. Equivalently, if the conditional probability (described in the text) of witnessing 'a' in a career when 'b' is present is large, then these two tokens are near one another according to our metric.

---

[11]Some OMA implementations allow for minor variations on these assignments; we use the version known as TDA written by Götz Rohwer (Blossfeld and Rohwer, 1999).

Note that the definition of $P(ab)$ requires a bit more attention. At a minimum, we must decide whether the joint probability is based on the entire sequence, or whether the estimation is limited to an observation window. $P(ab)$ refers to the probability of observing tokens a and b in the same sequence, but need they be near each other in the sequence? If an observation window is used, the metric may be further refined in at least two ways. The first allows heterogeneity across time, whereby $P_t(ab)$ is the probability of observing the pair in some interval beginning at t (ending, say within a year). The second enforces a direction to the observation: $P(ab)$ could be the probability of observing b within the year *following* a's occurrence.

We estimated a (time heterogeneous) version of $P_t(ab)$ that ignores the order of observation. $P_t(ab)$ is the probability of observing a and b in a two-year interval that begins at time t. This 500x500 matrix (for a given t) is interesting in its own right. We call it the "companion" matrix, for it identifies IxO (and the other OLF) tokens that have relationships with each other (and their strength). The diagonals of this matrix are one, since a token is always its own companion. For larger t, the off-diagonal probabilities tend to zero, as workers make fewer and fewer transitions overall.

Some tokens are extremely frequent while others are rare. To adjust for this, the original companion joint probability is often normalized by the marginal product of the token component frequencies. That is, if the probability of seeing A and B together is large compared to the product of their marginal probabilities (how often we would expect the pair by chance, under independence), then we have indication that A and B "belong" together in some manner. A version of this metric,

$$s(a,b) = \log\left(\frac{P(ab)}{P(a)P(b)}\right),$$

where s is the distance metric and a and b are tokens, is commonly applied in biological sequence analysis (Durbin, et al., 1998).

Implementing time heterogeneity with OMA

While it is reasonably easy to define a token similarity matrix based on the joint probability of companionship and the corresponding marginals, applying these to an OMA is a bit more involved. A most basic issue is the relationship between the similarity matrix and the insertion and deletion (shorthanded as 'indel') costs. The metric $s(a,b)$ spans the entire real line, while indel costs are fixed and finite. Our approach was to view $s(a,b)$ as a likelihood ratio, since two very simple models for sequences are implicitly being compared, ignoring all other tokens at each point of comparison. Referring the log likelihood ratio to a χ_1^2 distribution provided a distance between 0 and 1, which was easily rescaled so that indels would neither dominate nor become irrelevant.[12]

[12]Negative values of s were taken to be zero.

Time heterogeneity is more complicated, as many OMA implementations assume there is one set of substitution costs. Varying substitution cost within an OMA run may also be problematic, for the algorithm "realigns" the sequence through indels, so one may be assessing the cost of substituting a at time t for b at time t', which may be undefined. We implement heterogeneity by splitting the sequences into four segments and then running OMA on each separately, using a substitution cost matrix derived from $P_t(a,b)$ (we take an average), for all t in the segment being processed. We then add the distances from the four segments for an overall distance. This compromise still manages to exploit the "local" sequence alignments (minor noise or shifts) at which OMA excels.

This approach allows for a variety of adjustments to the distance computed by OMA. It implements a form of time heterogeneity, which is seen to be necessary in career data. The time-dependent distances can be reweighted to reflect the relative importance of different segments of the sequence. They can be aggregated non-additively — the minimum can be taken, for example. Lastly, this "divide and then recombine" approach can be extended to examine the very unit of analysis. We reran OMA on sequences whose tokens had been simplified to reflect industry or occupation only. Thus, we had a measure, for every pair, of the distance between them based on industry patterning, occupational patterning, or their joint patterning (we evaluated linear combinations of these to identify optimal clustering weights — results not discussed here). Other reasonable patterns to explore include education and training, and time out of the labor force.

Clustering dissimilarities

We make extensive use of the recursive partitioning algorithm known as PAM (partitioning around medoids; Kaufman and Rousseeuw, 1990). Unlike some clustering algorithms, these can take a distance matrix as input. This is extremely convenient, since we have no clear set of measures upon which to base our clusters, but OMA can provide a distance matrix that approximately reflects our four guiding principles (presented in the text) if we follow the procedures outlined above.

PAM algorithms often provide the useful diagnostic "silhouette width" (SW; a measure between -1 and 1) to assess how well items belong in the clusters to which they are assigned. Roughly speaking, SW is near zero or negative for a sequence where the nearest cluster medoid (center) is closer than the medoid of its assigned cluster.[13]
For substantive reasons, we form separate clusters for each mobility-based group (Stuck, Mobile, Never Low). The main parameter PAM requires is the number of clusters; we tried a variety of these and used the average silhouette width to guide our choice.

[13]The cluster assignments are globally optimal, meaning that moving any sequence out of that assignment reduces the overall goodness of fit. Locally, however, a sequence might appear to "belong" more to a nearby cluster even though moving it there would reduce the fit.

Displaying the career sequences

Displaying sequences for which there are 500 unique tokens is a challenge. We adopt the following strategy. The token components are from a smaller alphabet (of 25 industries and 20 occupations), so we assign symbol/color combinations to create these (see the legend at the end of Exhibit A). For example, waiters (eating and drinking industry / food service occupation) are assigned a pink triangle and black dot. The former, representing the industry is placed "above" the latter on the plot. To limit redundancy, an annual sampling (as opposed to quarterly) is displayed. So each worker has a sequence of up to 16 token (symbol) pairs. Non-working periods are filled in with codes 'E' for enrolled in school, 'O' for OLF and 'U' for unemployed. When this information is missing the point is left blank.

Further notes on presentation are worth mentioning. The medoid (see Kaufman and Rousseeuw, 1990) is given at the bottom of each cluster, providing some sense of a central theme. We write an abbreviated modal token for the medoid following "M=" under each plot. Here, other summaries are given, such as percent female, and race/ethnicity breakdowns. The cluster's silhouette width is given after "sw=" in the same area. The "drop=" segment will always be zero and can be ignored. Lastly, the sequences are sorted from those closest to the medoid to the furthest as we go from the bottom up. The raw distances are given in the rightmost shaded column, while the corresponding silhouette widths are in the leftmost.

Appendix C. Identifying matched pairs

In order to undertake the quasi-experimental comparison of similar clusters, we needed a method to determine the similarity of clusters. For each cluster in each mobility group, we wanted to determine its similarity to all the clusters in the other two mobility groups.

We borrowed the concept of "dissimilarity" that is often used in research on racial and other types of segregation. If one has two groups, the dissimilarity of the two groups is the percentage of individuals in one group that have to be exchanged with individuals in the other group in order to make the groups identical with respect to the property in question. For instance, if one group consists of 80 white people and 20 black people, and the other has 70 white people and 30 black people, the dissimilarity is 5 percent, because the groups can be made identical by exchanging 5 white people in the first group with 5 black people in the second group, making both groups have a 75/25 split.

Extending this to our analysis, we consider the IxO tokens that represent each of our clusters. We pool all of the IxO tokens in each cluster, from all the career sequences in the cluster, into a single set. We then compute, for each cluster, the share of that cluster represented by each token (this is marginal distribution of the tokens for the cluster). If p_{mci} represents the share of token *i* in cluster *c* of mobility group *m*, then the dissimilarity between two groups index by (m_1, c_1) and (m_2, c_2) is given by one half of the component-wise sum of the absolute value of the difference between the two vectors $\vec{p}_{m_1c_1} = (p_{m_1c_11}, \ldots, p_{m_1c_1K})'$ and $\vec{p}_{m_1c_1} = (p_{m_1c_11}, \ldots, p_{m_1c_1K})'$, where *K* is the number of unique IxO tokens. For the comparison, we looked only at industry-occupation pairs, excluding time spent out of the labor force. Thus, the dissimilarity

$$D(m_1c_1, m_2c_2) = \frac{1}{2}\sum_{i=1}^{K}\left|p_{m_1c_1i} - p_{m_2c_2i}\right|.$$

We computed all possible pair-wise dissimilarities between clusters across different mobility groups. There were (3×16×32)/2=768 such numbers. We examined the range of these dissimilarities graphically and determined that a cutoff value of 0.5 was appropriate for separating those pairs which were highly similar (with low dissimilarities) with the others.

This cutoff criterion generated fifteen similarity groupings, most of which were triples consisting of one Stuck, one Mobile, and one Never Low cluster, and a few of which were just Stuck and Mobile cluster pairs. In some cases, the triples were "tightly-similar" in that each cluster pair within the triple fell below the dissimilarity cutoff; in others, the Stuck and Mobile met the cutoff, as well as the Mobile and Never Low, but not the remaining pair. We did not include other combinations (such as Stuck being close to Never Low but not to Mobile), because such scenarios were irrelevant to our analysis.

From these triples and pairs, we were able to identify well-matched Stuck/Mobile cluster pairs for use in our quasi-experimental case studies.

References

Abbott, A. (1995). Sequence Analysis. *Annual Review of Sociology*, 21, 93-113.

Andersson, Fredrik, Holzer, Harry J., and Lane, Julia I. 2004. *Moving Up or Moving On: Who Advances in the Low-Wage Labor Market?* New York: Russell Sage Foundation.

Bernhardt, Annette, Morris, Martina, Handcock, Mark, and Scott, Marc. 2001. *Divergent Paths: Economic Mobility in the New American Labor Market*. New York: Russell Sage Foundation.

Blossfeld, H-P. and Rohwer, G. 1999. *Techniques of Event History Modeling*. Hillsdale, NJ: Lawrence Ehlbaum.

Boushey, Heather. 2005. No Way Out: How Prime-Age Workers Get Trapped in Minimum Wage Jobs. Briefing Paper. Washington, DC: Center for Economic and Policy Research.

Durbin, R., Eddy, S., Krogh, A., and Mitchison, G. 1998. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. New York: Cambridge University Press.

Gottschalk, Peter, and Moffitt, Robert. 1994. The Growth in Earnings Instability in the U.S. Labor Market. *Brookings Papers on Economic Activity,* 2: 217-54.

Hoff, P., Raftery, A. E., and Handcock, M. 2002. Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association*, 97, 1090-1098.

Kaufman, L. and Rousseeuw, P.J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.

Massey, Douglas S., and Denton, Nancy A. 1988. The dimensions of racial segregation. *Social Forces* 67: 281-315

Mitnik, Pablo A., and Zeidenberg, Matthew. 2004. Too Many Bad Jobs: An Analysis of the Prospects for Career Ladder Initiatives in the Service Economy. Paper presented at the 56th Annual Meeting of the Industrial Relations Research Association, San Diego, January.

Neckerman, Katherine (ed.). 2004. *Social Inequality*. New York, NY: Russell Sage.

Sankoff, D., and Kruskal, J. B. 1983. *Time Warps, String Edits, and Macromolecules*. Reading MA: Addison Wesley.

Spenner, Kenneth I., Luther B. Otto, Vaughn R.A. Call. 1982. *Career lines and careers*.

Lexington, Mass.: Lexington Books

Strawn, Julie, and Martinson, Karin. 2000. “Promoting Access to Better Jobs: Lessons for Job Advancement from Welfare Reform,” in *Low-Wage Workers in the New Economy*, edited by Richard Kazis and Marc S. Miller. Washington, D.C.: The Urban Institute Press.

Topel, Robert H., and Ward, Michael P., 1992. Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics*, vol. 107(2), 439-79.

Vingron, M., and Waterman, M. S. 1994. Sequence Alignment and Penalty Choice. *Journal of Molecular Biology*, 235:1-12.

Table 1. Variable Description

Variable name	Description
Wage growth, age 24 to 38 ($2002)	Mean permanent wage at age 24 and 38, in 2002 dollars
Cumulative work experience	Number of weeks worked, from age 16 to point of evaluation
% Time in dual jobs	Proportion of time a dual job is reported
% Public sector	Proportion of employment that is public sector
% Less than high school	Proportion of workers with less than a high school degree
% High school degree	Proportion of workers who have a high school degree, but no college
% En route Associate's	Percent who ever attained Associate's degree (but may have achieved higher)
% Some college	Proportion of workers have had some college, no four-year degree
% College degree	Proportion of workers with a four-year college degree
% Female	Proportion of female workers
% Black	Proportion of black workers
% Hispanic	Proportion of Latino workers
% Urban	Proportion of time urban (in SMSA)
% South	Proportion of time living in South

The measures below are reported in aggregate in Table 2, and for two age spans in Exhibit A: 20-25 and 31-36.

Variable name	Description
% Time spent out of labor force or unemployed (or both)	Proportion of time out of the labor force or unemployed or both (school enrollments excluded)
% Part-time work	Proportion of time work less than 35 hours
% Year-round work	Proportion of time work at least 40 weeks a year
# of employers	Number of distinct employers over a specified period
# of industries	Number of distinct two-digit industries over a specified period
# of occupations	Number of distinct two-digit occupations over a specified period
Average time spent in one industry-occupation category	Average length of run consisting of same job (industry-occupation pair), over a specified period
% Time in union	Proportion of time working a union job
% Employer provided training	Proportion of time in which one-month or more employer-provided training reported

Notes: All measures described as "percents" are based on quarterly jobs spanning ages 20-36. When we refer to a "job" we mean a distinct industry and occupation pair, such as retail department store sales clerk.

Table 2. Mean levels of key measures for three mobility groups

Note: All table entries are means except for the first and last rows.

Measure	Stuck	Mobile	Never Low	Overall
Share of workforce	*28.2*	*33.0*	*38.8*	*100.0*
Wage growth, age 24 to 38 ($2002)	$6.79 - $7.98	$8.31 - $15.38	$13.06 - $23.18	$9.72 - $16.32
Worker demographics:				
Percent female	68.2	52.0	35.4	50.1
Percent black	21.0	15.2	7.7	13.9
Percent Hispanic	6.8	7.0	5.6	6.4
Percent urban	69.8	77.5	86.7	78.9
Percent living in South	41.6	36.7	27.7	34.6
Educational attainment by age 36:				
Percent less than high school	20.6	10.3	4.8	11.1
Percent holding high school degree	52.9	38.2	32.2	40.0
Percent with some college (no degree)	19.1	24.0	22.4	22.0
Percent with 4-year college degree	7.4	27.4	40.6	26.9
Job characteristics:				
Proportion of public sector employment	7.7	13.6	11.7	11.2
Proportion of union employment	10.2	14.1	20.4	15.5
Employer-provided training (percent)	2.5	5.2	7.0	5.1
Career structure:				
Time OLF or unemployed (percent)	36.2	18.7	9.7	20.1
Time spent out of labor force (percent)	28.0	13.6	6.6	14.9
Time spent unemployed (percent)	8.1	4.9	3.1	5.1
Part-time work (percent)	28.4	19.1	14.2	19.8
Year-round work (percent)	52.1	70.3	81.3	69.4
Time holding dual jobs (percent)	7.3	8.4	8.5	8.1
Cumulative work experience (weeks)	585.8	722.7	810.1	718.0
Number of employers	7.0	7.1	5.8	6.6
Number of industries	4.7	4.8	4.1	4.5
Number of occupations	4.3	4.5	4.1	4.3
Number of observations	2484	2670	2558	7712

Table 3. Male/female comparisons

	Stuck		Mobile		Never Low	
Measure	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
Wage growth, age 24 to 38 ($2002)	$7.14 - $8.30	$6.63 - $7.84	$8.50 - 16.16	$8.14 - $14.69	$13.39 - $ 23.70	$12.45 - $22.18
Worker demographics:						
Percent black	27.4	18.1	15.6	14.8	7.4	8.2
Percent Hispanic	7.5	6.6	7.0	6.9	5.6	5.6
Percent urban	66.9	71.1	74.9	79.9	84.8	90.2
Percent living in South	43.0	41.0	36.6	36.8	27.4	28.2
Educational attainment by age 36:						
Percent less than high school	27.7	17.3	13.4	7.5	6.8	1.2
Percent holding high school degree	53.6	52.6	40.9	35.8	37.1	23.1
Percent with some college (no degree)	13.8	21.6	19.5	28.3	21.0	24.9
Percent who ever attained Associate’s	2.2	7.4	6.2	10.7	7.6	10.4
Percent with 4-year college degree	5.0	8.5	26.3	28.4	35.1	50.7
Job characteristics:						
Proportion of public sector employment	6.8	8.2	11.3	15.7	9.9	15.0
Proportion of union employment	10.7	10.0	14.5	13.8	22.2	17.0
Employer-provided training (percent)	1.8	2.8	5.2	5.2	6.9	7.1
Career structure:						
Time OLF or unemployed (percent)	26.3	40.8	13.4	23.5	7.4	13.9
Time spent out of labor force (percent)	14.2	34.4	7.1	19.6	3.8	11.7
Time spent unemployed (percent)	11.9	6.3	6.2	3.8	3.6	2.1
Part-time work (percent)	16.9	33.8	13.4	24.4	9.4	23.0

Table 4. Career clusters for three mobility groups

Stuck workers		Mobile workers		Never low workers	
Percent	Cluster description	Percent	Cluster description	Percent	Cluster description
18.0	Short, infrequent service jobs	17.6	Infrequent service jobs	10.2	Unionized non-durable manufacturing
13.9	Moderate-length stints in service jobs	9.5	Teachers and administrators	10.0	Union construction careers
9.4	Unemployed blue-collar workers	7.8	Craft construction workers	9.9	Computer programmers and engineers
8.6	Stable food service career	7.1	Retail hoppers	9.5	Unionized durable manufacturing
7.2	Late starters in retail and grocery sales	6.7	Durable manufacturing	8.6	Nurses, technicians, managers, doctors
6.7	Stable retail career	6.8	Flexible FIRE careers	7.9	Elite FIRE workers
6.2	Non-durable manufacturing	6.5	Civil servants	6.0	Accountants and lawyers
4.4	Health aides	5.6	Nurses, nursing aides, medical technicians	5.7	Auto mechanics, truck drivers
4.1	Durable manufacturing	5.5	Clerical careers in manufacturing/wholesale trade	5.7	Early labor market exiters
3.7	Agriculture workers	5.3	Unionized non-durable manufacturing	5.5	Retail managers and sales
3.6	Non-union construction and landscaping	5.2	Restaurant cooks, waiters & managers	4.8	Teachers and administrators
3.3	Early childhood educators	4.4	Truck and bus drivers	4.8	Postal service & other government
3.3	Secretaries and bank tellers	3.7	Auto mechanics	3.4	Unionized utilities
2.8	Auto mechanics and repairmen	3.3	Medical receptionists and secretaries	3.1	Restaurant managers and cooks
2.8	Part-time cleaners & laundry workers	3.0	Agriculture workers	2.5	Telecommunications careers
1.9	Hair dressers and childcare workers	2.2	Unionized grocery store workers	2.4	Police and firemen

Table 5. Matched career cluster pairs and triplets across mobility groups

Stuck workers		Mobile workers		Never low workers	
Percent	Cluster description	Percent	Cluster description	Percent	Cluster description
Strong relationship between industry-occupation pairs and patterning in Stuck and Mobile:					
3.7	Agriculture workers	3.0	Agriculture workers		
2.8	Auto mechanics and repairmen	3.7	Auto mechanics		
Strong relationship between industry-occupation pairs and patterning in Stuck, Mobile and Never Low:					
48.5	Short, infrequent service jobs; Moderate-length stints in service jobs; Unemployed blue-collar workers; Late starters in retail and grocery sales	17.6	Infrequent service jobs	5.7	Early labor market exiters
13.9	Stable retail career; Late starters in retail and grocery sales	7.1	Retail hoppers	5.5	Retail managers and sales
8.6	Stable food service career	5.2	Restaurant cooks, waiters & managers	3.1	Restaurant managers and cooks
6.2	Non-durable manufacturing	5.3	Unionized non-durable manufacturing	10.2	Unionized non-durable manufacturing
4.4	Health aides	5.6	Nurses, nursing aides, medical technicians	8.6	Nurses, technician, managers, doctors
4.1	Durable manufacturing	6.7	Durable manufacturing	9.5	Unionized durable manufacturing
3.6	Non-union construction and landscaping	7.8	Craft construction workers	10.0	Union construction careers
3.3	Secretaries and bank tellers	6.8	Flexible FIRE careers	7.9	Elite FIRE workers
3.3	Early childhood educators	9.5	Teachers and administrators	4.8	Teachers and administrators

NOTES: Several Stuck clusters matched with a single Mobile/Never Low pair, forming a complex triple with multiple "entry points." The "late starters in retail" Stuck cluster matches both service jobs and a retail cluster in the Mobile group. Percents reflect each cluster's share of the mobility group (or their sum).

Table 6. Comparing the Stuck and Mobile Health Care Clusters

Measure	Stuck Healthcare		Mobile Healthcare	
Cluster as percent of mobility group	4.4		5.6	
Wage growth, age 24 to 38 ($2002)	$6.91 – $8.47		$8.28 - $16.58	
Worker demographics:				
Percent female	82		75	
Percent black	28		22	
Percent Hispanic	8		5	
Percent urban	78		80	
Percent living in South	39		34	
Educational attainment by age 36:				
Percent less than high school	12		4	
Percent holding high school degree	51		33	
Percent with some college (no degree)	30		37	
Percent with 4-year college degree	7		26	
Job characteristics:				
Cumulative work experience (weeks)	679		766	
Time holding dual jobs (percent)	9		9	
Proportion of public sector employment	12		15	
Career structure:	*Early 20s*	*Mid-30s*	*Early 20s*	*Mid-30s*
Time OLF or unemployed (percent)	41	10	24	7
Time spent out of labor force (percent)	31	6	17	6
Time spent unemployed (percent)	10	4	6	1
Part-time work (percent)	20	23	38	19
Year-round work (percent)	43	84	59	91
Number of employers	3.1	2.6	3.2	2.1
Number of industries	2.7	2.0	2.5	1.5
Number of occupations	2.6	2.4	3.1	2.0
Average time spent in one industry-occupation category (quarters)	2.2	5.4	2.6	6.1
Proportion of union employment	17	14	15	21
Employer-provided training (percent)	0	6	1	9

Table 7. Comparing the Stuck and Mobile Durable Manufacturing Clusters

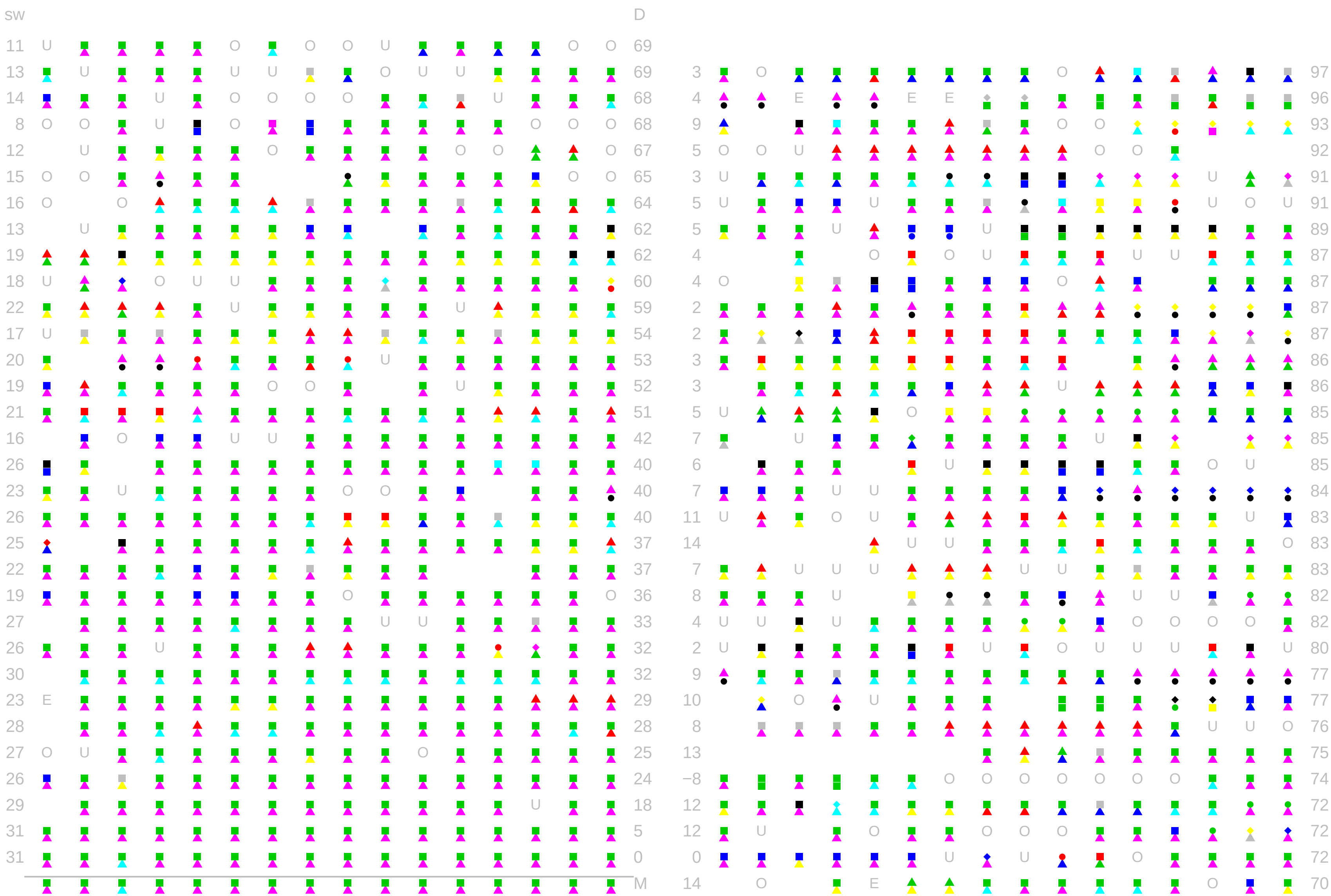
Measure	Stuck Durable Manufacturing		Mobile Durable Manufacturing	
Cluster as percent of mobility group	4.1		6.7	
Wage growth, age 24 to 38 ($2002)	$7.73 - $8.97		$8.31 - $15.38	
Worker demographics:				
Percent female	40		22	
Percent black	17		10	
Percent Hispanic	5		7	
Percent urban	60		69	
Percent living in South	50		29	
Educational attainment by age 36:				
Percent less than high school	23		19	
Percent holding high school degree	62		62	
Percent with some college (no degree)	16		15	
Percent with 4-year college degree	0		5	
Job characteristics:				
Cumulative work experience (weeks)	788		784	
Time holding dual jobs (percent)	5		7	
Proportion of public sector employment	2		3	
Career structure:	*Early 20s*	*Mid-30s*	*Early 20s*	*Mid-30s*
Time OLF or unemployed (percent)	17	13	16	5
Time spent out of labor force (percent)	10	7	8	3
Time spent unemployed (percent)	7	6	8	3
Part-time work (percent)	9	11	16	6
Year-round work (percent)	71	82	66	91
Number of employers	3.0	2.4	3.5	2.0
Number of industries	2.5	2.2	2.8	1.9
Number of occupations	2.6	2.5	2.8	2.1
Average time spent in one industry-occupation category (quarters)	2.6	3.9	2.3	5.2
Proportion of union employment	20	11	14	23
Employer-provided training (percent)	2	4	1	8

Exhibit A: Career Sequences

Matched pairs in durable and non-durable manufacturing

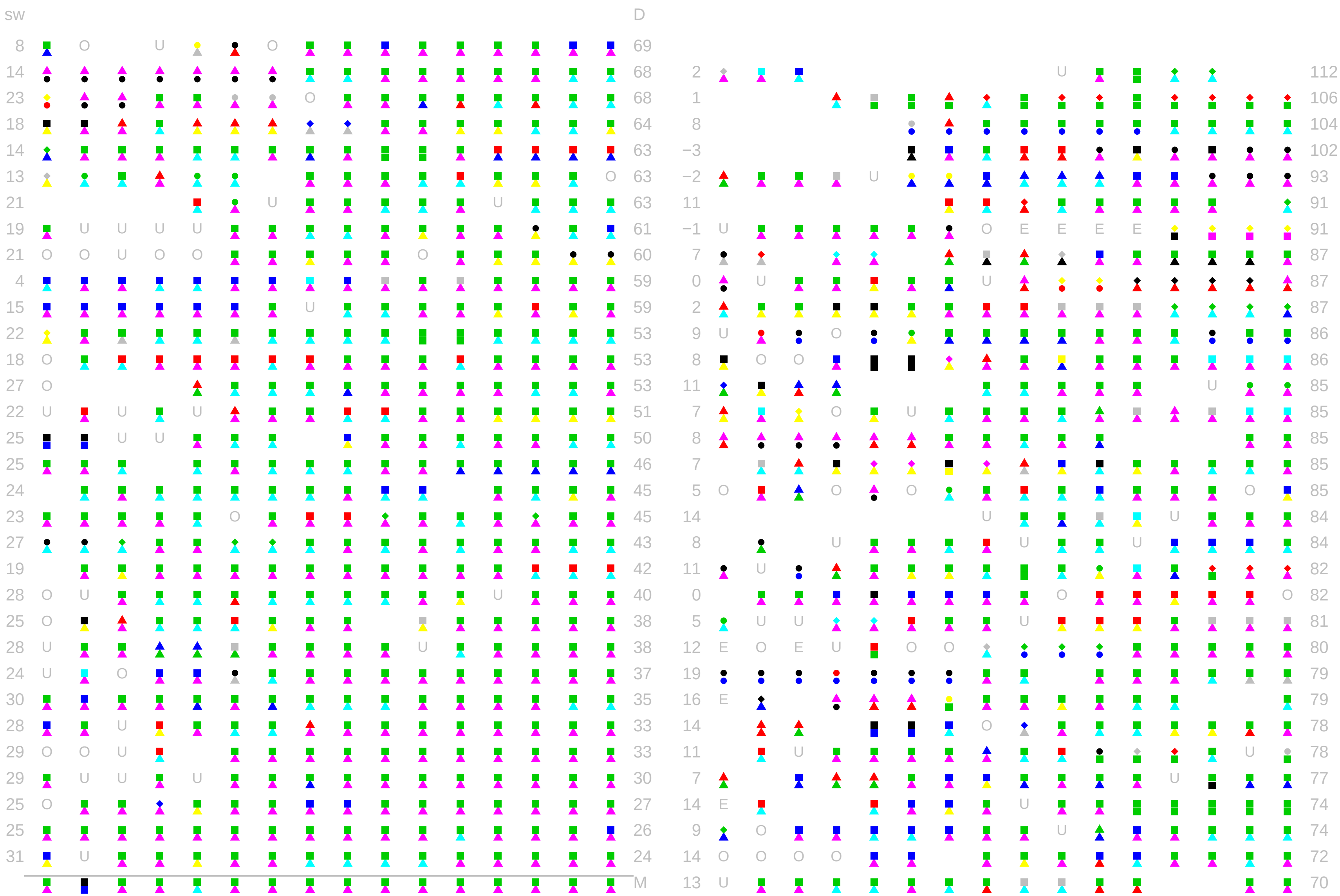
- Each row represents an individual career sequence.
- Each row has 16 observations, representing the employment status of the individual at a given time in one calendar year. (The full sequences are actually made up of 64 quarters, but for illustration, we show only show the second quarter from each year here).
- Each observation represents either a job, or the individual's labor force status (unemployed, out of labor force, enrolled in school). The jobs are described in terms of industry (the top symbol) and occupation (the bottom symbol).
- The left panel tends to show the dominant career path in that cluster, and the right panel tends to show variations on that dominant path.
- The symbol legend is attached at the end.

stuck workers, IxO−based clusters; cluster 6



f=4%; n=89; sw=0.12 (0.14); drop=0% M=mnfD:Oper
%Fem=40 (wht,blk,hsp)=(78,17,5)%

mobile workers, IxO−based clusters; cluster 5



f=7%; n=170; sw=0.14 (0.15); drop=0% M=mnfD:Oper
%Fem=22 (wht,blk,hsp)=(83,10,7)%

Industry

1	agriculture,forestry,fishing,mining
2	construction
3	manufacturing durable
4	manufacturing non−durable
5	transportation
6	communications
7	utilities,sanitary services
8	wholesale trade
9	retail hard goods (except auto)
10	retail food
11	retail auto
12	eating & drinking
13	finance,insurance,real estate
14	high end business services
15	building services
16	temp agencies
17	auto & repair services
18	hotels & laundry
19	personal services
20	entertainment,recreation services
21	healthcare
22	education
23	non−profit
24	professional services
25	public administration

Occupation

1	professionals (drs., lawyers, etc.)
2	engineers, scientists, non−health tech
3	farm laborers, farm foremen
4	nurses,dietician,therapist
5	health technologists,technicians
6	religious,social sci,social wrk
7	teachers
8	writers,artists,entertainers
9	managers, admin
10	sales wrkrs
11	clerical,unskilled wrkrs
12	craft workers and mechanics
13	operatives
14	laborers except farm
15	cleaning service wrkrs
16	food service wrkrs
17	health service wrkrs
18	personal service wrkrs
19	protective service wrkrs
20	private household wrkrs

Other

E	Enrl
O	OLF
U	UNEMP
	Msg IxO or Unknown

Abbreviations:
sw=silhouette width
D=distance from medoid
M=medoid
Note: sw value in parentheses is average sw after cleaning

Exhibit B: Detailed cluster profiles

NOTES: The structure of this data is rich enough to warrant detailed descriptions of each career cluster that we identify. The same summary measures used in the figures are computed for each cluster. Instead of reporting means at the cluster level, we report the tertile to which they belong: lower, middle, or upper third. This describes whether a feature is unusually low, typical, or high (abbreviated ‘lo’ ‘med’, ‘hi’). It is important to avoid interpreting these in any absolute sense. For example, Latinos account for 6.4% of the analysis sample, so we report that “% Hispanic” is ‘hi’ in a cluster that is 10% Latino, but these still comprise a small share. In addition to aiding between-cluster comparisons, this approach adjusts for secular trends, such as decline in union membership over time: union rates of 20% for 20-25 year olds may be in the same tertile as rates of 10% are for 31-36 year olds ten years later. Four detailed cluster summaries for the two matched pair case studies and four additional example clusters are given below. A graphical representation of the raw career sequence data was given in Exhibit A, and the cluster numbers on those graphics provide a link to a corresponding cluster profile.

Stuck Cluster #2: Secretaries and bank tellers

Overview: Almost entirely female cluster (98%). Most workers have either high school degree or some college experience. Good amount of training, especially early on. Little union membership.

Dominant career path: Secretaries, bank tellers, and other clerical occupations, largely in banking and insurance. Usually direct entry into these jobs, though some are coming from retail sales works. At first, strong labor force attachment and low rates of job changing. Later on, see more dropping out of labor market and more part-time jobs (possibly to take care of children). Mild evidence of mobility: managerial occupations increase from 7% to 15% over time.

Variations on dominant path: Workers stay in clerical occupations, but move between a number of different industries (retail, hotels, manufacturing, health care, construction, professional services). Less evidence of occupational mobility, and more time spent out of labor force.

Summary variables	
Cluster as percent of stuck group	3.3
Wage growth, age 24 to 38 ($2002)	$7.83 - $8.38
Cumulative work experience	Med
% Time in dual jobs	Med
% Public sector	Med
% Less than high school	Lo
% High school degree	Hi
% Some college	Hi
% College degree	Med
% Female	Hi
% Black	Lo
% Hispanic	Lo
% Urban	Lo
% South	Med

Changes over career		
	Mid-20s	*Late 30s*
% Time spent out of labor force or unemployed	Lo	Hi
% Time spent OLF	Med	Hi
% Time spent unemployed	Lo	Med
% Part-time work	Med	Hi
% Year-round work	Hi	Lo
# of employers	Lo	Med
# of industries	Lo	Hi
# of occupations	Lo	Med
Average time spent in one industry-occupation category	Hi	Med
% Time in union	Lo	Lo
% Time received training	Hi	Med

Details on industries and occupations	Percent of jobs
Top three jobs, mid-20s	
1. Industry: FIRE, largely banking and insurance Occupation: Secretaries, bank tellers, bookkeepers, and other clerical workers	42.7
2. Industry: Retail non-food, drug stores, and direct selling establishments Occupation: Sales clerks, sales managers, cashiers	6.4
3. Industry: FIRE, largely banking and insurance Occupation: Bank managers, office managers	5.8
Top three jobs, late 30s	
1. Industry: FIRE, largely banking and insurance Occupation: Secretaries, bank tellers, bookkeepers, and other clerical workers	42.9
2. Industry: FIRE, largely banking and insurance Occupation: Bank officers, managers, and financial managers	8.3
3. Industry: Health care, largely hospitals and nursing homes Occupation: Secretaries and other clerical workers	3.0

Stuck Cluster #6: Durable manufacturing

Overview: More than half of workers have high school degree (**62**%), but a quarter are drop-outs. Higher representation of black workers. Little training throughout. Unionization is high early on at 20%, but then declines to 11% ten years later.

Dominant career path: Factory workers in durable manufacturing (furniture, sawmills, cars). Assemblers, machine operators, forklift or press operators, etc. Most workers directly enter the industry and show sustained work in the industry, despite some unemployment spells scattered throughout the career.

Variations on dominant path: Much more unemployment (likely layoffs), followed by considerable job changing across industries and occupations, especially by mid-30s. Common changes are to retail jobs (including gas service stations and lumberyard work); to non-durable manufacturing; or to construction, wholesale trade or food service work. Occupations are still largely blue collar.

Summary variables	
Cluster as percent of stuck group	4.1
Wage growth, age 24 to 38 ($2002)	$7.73 - $8.97
Cumulative work experience	Med
% Time in dual jobs	Lo
% Public sector	Lo
% Less than high school	Hi
% High school degree	Hi
% Some college	Lo
% College degree	Lo
% Female	Med
% Black	Med
% Hispanic	Lo
% Urban	Lo
% South	Hi

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Med	Hi
% Time spent OLF	Med	Hi
% Time spent unemployed	Med	Hi
% Part-time work	Lo	Med
% Year-round work	Med	Lo
# of employers	Lo	Med
# of industries	Lo	Hi
# of occupations	Lo	Hi
Average time spent in one	Med	Lo
industry-occupation category		
% Time in union	Hi	Med
% Time received training	Lo	Lo

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Durable manufacturing (furniture, cars, millwork, electrical machinery, glass, metal) Occupation: Assemblers, operatives (machine, forklift, press), welders, inspectors	38.7
2. Industry: Durable manufacturing (furniture, cars, machinery, millwork, metal) Occupation: Inspectors, machinists, foremen, woodworkers	8.5
3. Industry: Durable manufacturing (millwork, electrical machinery, logging, cars) Occupation: Freight handlers, lumbermen, misc. laborers	6.7
Top three jobs, mid-30s	
1. Industry: Durable manufacturing (furniture, cars, millwork, electrical machinery, glass, metal)) Occupation: Assemblers, operatives (machine, forklift, press), welders, inspectors	38.2
2. Industry: Durable manufacturing (furniture, cars, machinery, millwork, metal) Occupation: Inspectors, machinists, foremen, woodworkers	8.6
3. Industry: Durable manufacturing (millwork, electrical machinery, logging, cars) Occupation: Freight handlers, lumbermen, misc. laborers	8.0

Stuck Cluster #10: Health aides

Overview: Female dominated, with black workers over-represented. High rates of unemployment, especially early on. The majority (81%) have a high school degree or some college experience.

Dominant career path: Delayed workforce entry, followed by steady employment in nursing homes and hospitals, as health aides. Some switching to technicians, clerks, and practical nurse jobs, but in general little evidence of mobility.

Variations on dominant path: Lots of employer turnover and occupation changing, particularly later in career. Health aid jobs still form the bulk of the jobs held, but workers also mix in stints as cashiers or receptionists, primarily in eating and drinking places, laundries and transportation.

Summary variables	
Cluster as percent of stuck group	4.4
Wage growth, age 24 to 38 ($2002)	$6.91 - $8.47
Cumulative work experience	Lo
% Time in dual jobs	Med
% Public sector	Hi
% Less than high school	Med
% High school degree	Med
% Some college	Hi
% College degree	Med
% Female	Hi
% Black	Hi
% Hispanic	Hi
% Urban	Med
% South	Hi

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Hi	Med
% Time spent OLF	Hi	Med
% Time spent unemployed	Hi	Hi
% Part-time work	Med	Hi
% Year-round work	Lo	Lo
# of employers	Lo	Hi
# of industries	Lo	Med
# of occupations	Lo	Hi
Average time spent in one	Med	Med
industry-occupation category		
% Time in union	Hi	Med
% Time received training	Lo	Med

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Nursing homes and hospitals Occupation: Nursing/health aide, orderlies, attendant	23.8
2. Industry: eating & drinking Occupation: Food service workers, cooks, waiters, dishwashers	10.7
3. Industry: Nursing homes and hospitals Occupation: billing clerks, receptionists, secretaries	6.6
Top three jobs, mid-30s	
1. Industry: Nursing homes and hospitals Occupation: Nursing/health aide, orderlies, attendant	37.9
2. Industry: Nursing homes and hospitals Occupation: billing clerks, receptionists, secretaries	10.9
3. Industry: Nursing homes and hospitals Occupation: Health technologists and technicians	5.7

Mobile Cluster #5: Durable manufacturing

Overview: Male dominated cluster. More than half of workers have high school degree (**62**%), but **19**% are drop-outs. Moderate training throughout, and unionization is high at 23% by mid-30s.

Dominant career path: Operators, foremen and craft workers in durable manufacturing. Workers either enter industry directly, or by first passing through non-durable manufacturing. Labor force participation is sustained, usually with full-time, year-round work. Movement back and forth between operative and craft occupations is common. The percent of foremen doubles in a 10 year period, suggesting some job mobility.

Variations on dominant path: Greater unemployment (likely layoffs) and dropping out of labor force. While these trajectories are still clearly in durable manufacturing, interruptions lead to more varied occupations in the industry, as well as brief stints other industries (e.g. retail).

Summary variables	
Cluster as percent of mobile group	6.7
Wage growth, age 24 to 38 ($2002)	$8.31 - $15.38
Cumulative work experience	Med
% Time in dual jobs	Lo
% Public sector	Lo
% Less than high school	Hi
% High school degree	Hi
% Some college	Lo
% College degree	Lo
% Female	Lo
% Black	Med
% Hispanic	Med
% Urban	Lo
% South	Lo

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Med	Lo
% Time spent OLF	Med	Lo
% Time spent unemployed	Hi	Med
% Part-time work	Lo	Lo
% Year-round work	Med	Med
# of employers	Med	Med
# of industries	Med	Med
# of occupations	Med	Med
Average time spent in one industry-occupation category	Med	Med
% Time in union	Med	Hi
% Time received training	Lo	Med

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Durable manufacturing (cars, furniture, machinery, metal products) Occupation: Assemblers, machine operatives, welders, truck drivers	22.9
2. Industry: Durable manufacturing (cars, furniture, machinery) Occupation: Foremen, machinists, cabinetmakers, mechanics, crane operators	8.6
3. Industry: Non-durable manufacturing (unspecified, apparel, plastics, meat) Occupation: operatives (machine, forklift, press), welders, assemblers	6.9
Top three jobs, mid-30s	
1. Industry: Durable manufacturing (cars, furniture, machinery, metal products) Occupation: Assemblers, machine operatives, welders, truck drivers	39.6
2. Industry: Durable manufacturing (cars, furniture, machinery) Occupation: Foremen, machinists, cabinetmakers, mechanics, crane operators	16.2
3. Industry: Durable manufacturing (logging, cars, construction supplies, furniture, millwork) Occupation: Misc. laborers, lumbermen, freight handlers	6.3

Mobile Cluster #11: Nurses, nursing aides, medical technicians

Overview: Disproportionately female and black cluster. About 65% have some college experience or four-year college degree. Notably, 20% attain an Associate's degree at some point. Part-time work and job shopping is high early on. Union membership tends to be higher, but training takes place later in the career.

Dominant career path: Nurses (both RNs and LPNs), nursing aides, and medical technicians working in hospitals and nursing homes. Largely separate occupational tracks, but there is some evidence of mobility from nursing aide positions to either nurses or medical technician jobs. Delayed entry into labor market, then followed by sustained attachment (with some spells out of labor market, likely for childrearing). By the mid-30s, very little occupational changing.

Variations on dominant path: Significantly more time out of the labor market. While nursing is clear endpoint for these careers, they begin with a much wider range of jobs, including clerical and food service work; the shift to nursing doesn't happen until late 20s.

Summary variables	
Cluster as percent of mobile group	5.6
Wage growth, age 24 to 38 ($2002)	$8.28 - 16.58
Cumulative work experience	Lo
% Time in dual jobs	Med
% Public sector	Hi
% Less than high school	Med
% High school degree	Lo
% Some college	Hi
% College degree	Med
% Female	Hi
% Black	Hi
% Hispanic	Lo
% Urban	Med
% South	Med

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Hi	Med
% Time spent OLF	Hi	Med
% Time spent unemployed	Med	Lo
% Part-time work	Hi	Med
% Year-round work	Lo	Med
# of employers	Lo	Med
# of industries	Lo	Lo
# of occupations	Hi	Lo
Average time spent in one industry-occupation category	Med	Hi
% Time in union	Med	Hi
% Time received training	Lo	Hi

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Nursing homes, hospitals, dentist's offices Occupation: Nursing/health aides, orderlies, attendants, dental assistants	26.0
2. Industry: Hospitals, doctor/dentist offices Occupation: Health technologists and technicians, dental hygienists	8.3
3. Industry: Eating and drinking places Occupation: Waiters, food counter & service workers, cooks	6.4
Top three jobs, mid-30s	
1. Industry: Hospitals, nursing homes Occupation: Registered and practical nurses, therapists	26.1
2. Industry: Nursing homes, hospitals, dentist's offices Occupation: Nursing/health aides, orderlies, attendants, dental assistants	22.9
3. Industry: Hospitals, doctor's offices Occupation: Health technologists and technicians, dental hygienists	11.7

Mobile Cluster #12: Retail hoppers

Overview: Most workers have some college experience, or college degrees. Modest training early on, but then becomes high by mid-30s. Little union membership throughout.

Dominant career path: Sales clerks, department heads, and managers in a range of retail segments (department stores, lumber and hardware, apparel stores, drug stores). Workers finish education and then jump directly into the industry. Labor force participation is high, with frequent movement across employers and occupations throughout the career. Strong evidence of mobility: managerial occupations increase from about 10% to 20% over time. Many managers start in entry-level jobs.

Variations on dominant path: Workers finish education and jump into retail, but then by mid-30s increasingly move into other industries, sometimes as managers (wholesale trade, building services, professional services, manufacturing). There are more interruptions to labor force participation later in career.

Summary variables	
Cluster as percent of stuck group	7.1
Wage growth, age 24 to 38 ($2002)	$8.32 - 17.71
Cumulative work experience	Hi
% Time in dual jobs	Hi
% Public sector	Med
% Less than high school	Med
% High school degree	Lo
% Some college	Hi
% College degree	Hi
% Female	Med
% Black	Lo
% Hispanic	Hi
% Urban	Hi
% South	Lo

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Lo	Lo
% Time spent OLF	Med	Med
% Time spent unemployed	Med	Med
% Part-time work	Med	Med
% Year-round work	Med	Med
# of employers	Hi	Hi
# of industries	Hi	Hi
# of occupations	Hi	Hi
Average time spent in one	Med	Lo
industry-occupation category		
% Time in union	Lo	Lo
% Time received training	Hi	Hi

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Sales clerks, sales managers and department heads	32.3
2. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Managers and administrators	7.3
3. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Bookkeepers, stock clerks, clerical, shipping and receiving	7.0
Top three jobs, mid-30s	
1. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Sales clerks, sales managers and department heads	24.9
2. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Managers and administrators	9.4
3. Industry: Retail, largely department stores, misc. retail, lumber & hardware, apparel stores Occupation: Bookkeepers, stock clerks, clerical, shipping and receiving	5.9

Never Low Cluster #9: Restaurant managers and cooks

Overview: Most workers have either some college experience or high school degree. Women and black workers are under-represented. Moderately high union membership (likely in hotels or corporate cafeterias) but little employer-provided training early in career.

Dominant career path: Managers and food service workers (especially cooks) in eating and drinking places and hotel restaurants. After education, most workers enter the industry directly, followed by sustained labor force participation. High rates of job changing, but largely within the industry (though there is some movement to and from hotel restaurants). Strong evidence of mobility: managerial occupations increase from about 15% to 35% over time (as does training).

Variations on dominant path: More frequent dropping out of labor force. As well, workers mix restaurant jobs with jobs in recreation and entertainment and retail, often as managers. There appears to be a distinct path from restaurant work to airline attendant jobs.

Summary variables	
Cluster as percent of stuck group	3.1
Wage growth, age 24 to 38 ($2002)	$12.44 - 18.69
Cumulative work experience	Hi
% Time in dual jobs	Hi
% Public sector	Lo
% Less than high school	Med
% High school degree	Lo
% Some college	Hi
% College degree	Hi
% Female	Lo
% Black	Lo
% Hispanic	Med
% Urban	Hi
% South	Lo

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Lo	Med
% Time spent OLF	Med	Med
% Time spent unemployed	Lo	Med
% Part-time work	Med	Med
% Year-round work	Hi	Med
# of employers	Hi	Hi
# of industries	Hi	Med
# of occupations	Hi	Med
Average time spent in one	Med	Med
industry-occupation category		
% Time in union	Med	Med
% Time received training	Lo	Med

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Eating and drinking places Occupation: Cooks, waiters, bartenders	30.5
2. Industry: Eating and drinking places Occupation: Restaurant, cafeteria, and bar managers	13.4
3. Industry: Hotels and motels Occupation: Cooks, waiters, bartenders	6.5
Top three jobs, mid-30s	
1. Industry: Eating and drinking places Occupation: Restaurant, cafeteria, and bar managers	23.0
2. Industry: Eating and drinking places Occupation: Cooks, waiters, bartenders	18.8
3. Industry: Hotels and motels Occupation: Cooks, waiters, bartenders	8.4

Never Low Cluster #16: Auto mechanics, truck drivers

Overview: Moderately educated, almost exclusively male group. Significant training and unionization rates throughout the career. Workers enter the labor market either directly after schooling, or after a moderate spell of unemployment. Once established, these careers show low rates of job changing.

Dominant career path: These careers come in three flavors: auto service work at dealerships, auto service work at repair shops, and a largely separate trajectory for truck and train driver jobs. Careers consist of one primary job, with brief shopping. Overall, reasonable opportunity for advancement exists.

Variations on dominant path: Delayed entry into the labor force because of unemployment and education. Auto service jobs are more of a bridge to truck and train driver jobs. Managerial occupations in transportation and auto industries comprise about 15% of jobs by mid-30s.

Summary variables	
Cluster as percent of stuck group	5.7
Wage growth, age 24 to 38 ($2002)	$13.13 - $20.37
Cumulative work experience	Hi
% Time in dual jobs	Med
% Public sector	Med
% Less than high school	Med
% High school degree	Med
% Some college	Med
% College degree	Med
% Female	Lo
% Black	Lo
% Hispanic	Med
% Urban	Med
% South	Med

Changes over career		
	Early 20s	*Mid-30s*
% Time spent out of labor force or unemployed	Lo	Lo
% Time spent OLF	Lo	Lo
% Time spent unemployed	Lo	Lo
% Part-time work	Lo	Lo
% Year-round work	Hi	Hi
# of employers	Med	Med
# of industries	Med	Lo
# of occupations	Med	Lo
Average time spent in one	Hi	Hi
industry-occupation category		
% Time in union	Hi	Hi
% Time received training	Med	Hi

Details on industries and occupations	Percent of jobs
Top three jobs, early 20s	
1. Industry: Automobile services and repair services. Occupation: Auto mechanics, repairmen	9.3
2. Industry: Motor vehicle dealers Occupation: Auto mechanics, repairmen	7.6
3. Industry: Trucking, air, street railway and bus services Occupation: Truck drivers, delivery & routemen	7.5
Top three jobs, mid-30s	
1. Industry: Trucking, air, street railway and bus services Occupation: Truck drivers, delivery & routemen	16.7
2. Industry: Trucking, air, street railway and bus services Occupation: Manager, administrator, pilot, conductor	10.3
3. Industry: Automobile services and repair services. Occupation: Auto mechanics, repairmen	10.2