1-1-2017

The New Hires Quality Index: A Wage Metric for Newly Hired Workers

Brad J. Hershbein
W.E. Upjohn Institute for Employment Research, hershbein@upjohn.org

Citation

Hershbein, Brad J. 2017. "The New Hires Quality Index: A Wage Metric for Newly Hired Workers." Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
https://research.upjohn.org/reports/228

The New Hires Quality Index:
A Wage Metric for Newly Hired Workers

Brad Hershbein
W.E. Upjohn Institute for Employment Research
hershbein@upjohn.org

ABSTRACT

There is substantial interest in measuring not just the quantity of new jobs but the quality as well. Existing surveys by the Bureau of Labor Statistics describe the number of new jobs created each month, as well as wages of incumbent workers, but not wages (or other characteristics) of newly created jobs. This paper aims to fill that gap by describing the creation of a new job quality metric for newly hired workers. Drawing on publicly accessible data from the Current Population Survey and Occupational Employment Statistics, I construct a job quality index of new hires based on occupation, which is more closely tied to skill demand and wages than measures based on industry. The New Hires Quality Index accounts for changing demographics of hires, can be consistently constructed from 2001 forward overall and for subgroups, does not rely on self-reported wages, and can be updated monthly. It also permits decompositions to illustrate which groups are affecting changes in the index. It should serve as a valuable new tool in gauging realized labor demand.

Key words: new hires, job quality, wage index, Occupational Employment Statistics, Current Population Survey

I. Introduction

The statistical agencies of the federal government frequently release data on the creation of new jobs. Each month, for example, the Bureau of Labor Statistics (BLS) reports on the number of new jobs that employers have created, and these numbers are widely reported in the media and used by policymakers, such as the Federal Reserve, in shaping economic projections. Although these jobs numbers are timely and considered to be high quality because of large sample sizes, there is little information about them other than the industries in which they were created. Notably, details are missing about the demographics of workers taking the new jobs, or what occupations they constitute, or even what they pay. While BLS does collect and report information on wages in the jobs survey, the statistics are for *all* incumbent workers, not new hires. As a result, we know little about the quality of new jobs being created each month.

Yet, understanding characteristics of new jobs—and the workers in them—is a key policy interest. Are the jobs being created low-paying or high-paying? If many jobs were created but they were disproportionately low-paying, is that better or worse than if fewer jobs were created but they were higher-paying? Are the new jobs going to a broad cross-section of the population, or are they concentrated among certain groups? The answers to these questions could act as an important contemporaneous, and perhaps leading, indicator of economic activity. They could provide new insight into labor markets during downturns, recoveries, and booms. And they could shed light on deeper, more structural changes in skill demand as automation becomes more common in the workplace.

This paper describes the creation of a new index designed to measure the earnings power of recently hired workers and help answer the above questions. Drawing on the Current Population Survey, the same data source used to construct the official unemployment rate and labor force statistics, the index is based on the occupations—not industries—of newly hired workers. These detailed occupations of new hires are merged with wage data by occupation from a separate BLS survey, the Occupational Employment Statistics, and wages are then statistically adjusted for the demographics of newly hired workers. The result is a measure of the quality of new job hires—as captured by adjusted occupational wages—that can be calculated each month: the New Hires Quality Index (NHQI).

In terms of capturing job pay or "quality," the distinction between occupation and industry is important. An occupation describes what the worker does, capturing the skills used to perform various tasks. An industry describes the product or service made by the worker's employer. A worker's occupation provides a much better indication of pay than the worker's industry. This should intuitively make sense. A surgeon and a medical assistant often both work in the same industry—hospitals, within the broader industry of health care—but they do vastly different things, and a surgeon is often paid many times what a medical assistant is paid. Or, to take a different tack, a janitor may work for a restaurant or for an investment bank—two very different industries, with rather large differences in average wages—but the janitor's pay is still going to reflect what work is being done rather than where it's being done.[1] Indeed, a worker's occupation explains four times the variation in hourly wages as a worker's industry.[2] Thus, to understand the skills that employers want and the characteristics of newly realized hires, knowing occupation is better than knowing merely industry.

Furthermore, because the NHQI uses occupational data drawn from newly hired individuals, the index can be calculated for all new hires as well as for various worker subgroups, such as by sex, age, education, and other characteristics. Moreover, the volume of new hires each month can also be calculated for all these groups and analyzed in conjunction with the wage component.

The NHQI has a reasonably long time series and can be updated in a timely fashion. It is available for each month back to January 2001 and can be updated with approximately a four-week lag. (Thus, the NHQI for August 2017 can be created by the end of September 2017.) It is also easily and freely accessible. An interactive charting feature allowing the user to plot and compare the data across groups, as well as download subsets of the data, can be found at www.upjohn.org/nhqi.

The NHQI is designed to be a versatile tool, but it is crucial to understand what it is and what it is not. It **is** a measure of the realized demand for skill (i.e., "quality") of newly hired workers, as proxied by the occupations those workers are hired into. Since higher-skill (i.e., "higher-quality") jobs tend to pay more, the NHQI will rise when the occupation mix of new hires shifts to higher-paying occupations, and it will fall when the occupational mix of new hires shifts to lower-paying occupations. The NHQI is **not** a measure of actual wages of new hires. As such, it cannot measure how starting wages are changing *within* an occupation. Put differently, the NHQI will not capture rising pay over time for, say, software engineers, or distinguish the hiring of a brilliant software engineer from a merely good one. It does, however, track at a high level whether new hires are taking higher- or lower-paying occupations.[3] Consequently, it will no longer be necessary to try to infer whether "good" or "bad" jobs are being created based on industry-level job counts; the NHQI provides a more accurate and more direct measure of the hiring taking place.

This report describes the methodology used to construct the NHQI. Because the creation of any index is subject to some methodological choices, the report also explores the sensitivity of the NHQI to certain alternative choices. Additionally, there is a comparison of the NHQI with a

[1] The Occupational Employment Statistics show that in 2016 building cleaning workers in the finance industry earn $11.25 per hour at the median, compared with $10.61 per hour for janitors at restaurants. The median hourly wages for all workers in these industries are $24.41 and $9.80, respectively.

[2] This follows from a regression of log inflation-adjusted hourly wages on three-digit Standard Occupational Classification (SOC) and three-digit North American Industrial Classification System (NAICS) codes in the CPS. Even after adding additional covariates, such as metro status, age, sex, race/ethnicity, education, and nativity, the occupation codes have two-and-one-half times the explanatory power of industry codes, and about twice that of education.

[3] A comparison of the NHQI with a measure tracking the actual wages of new hires is presented in Section V of this report. That section also describes some of the pitfalls of the latter measure.

measure tracking actual reported wages of new hires; although this latter approach sounds appealing, a discussion details several difficulties with it, and why an occupational-based measure is more broadly useful. The NHQI should serve as a valuable indicator of the quality of new job hires for policymakers, journalists, businesses, researchers, and anyone else interested in job creation.

II. Methodology

Construction of the NHQI draws on two major, publicly accessible data sources. The first is the Current Population Survey (CPS), produced jointly by BLS and the Census Bureau. The CPS is a large-scale survey of approximately 60,000 households each month and is the nation's primary source for employment and unemployment statistics. For example, it is used to produce the headline monthly unemployment rate. It is also partially longitudinal in nature, with households interviewed several times over a span of up to 16 months. This feature, along with the availability of individual-level microdata on a frequent basis and questions about switching jobs, is crucial to constructing the index.

The second major data source is the Occupational Employment Statistics (OES), released annually by the Bureau of Labor Statistics. The OES asks businesses to report the number of their employees by detailed occupation, as well as the distribution of wages of each detailed occupation. Although microdata at the establishment level are not released publicly, aggregated statistics show the wage distribution for specific occupations at different levels of geography and industry. Importantly, these wage data can be matched to the occupations that people take in the CPS, allowing consistent wage information.

A. Current Population Survey

A.1 Matching Individuals from Month to Month

Households are interviewed for the CPS in a rotating panel. Specifically, a household is scheduled to be interviewed for four months in a row, be off for eight months, and then interviewed again during each of the next four months. A household's month in the panel is called its rotation group, which ranges from one to eight. Each month one-eighth of the total sample is added as a new sample household (rotation group 1) to replace the one-eighth that exits (rotation group 8 in the previous month). Thus, households in rotation groups 1, 2, and 3 will be sampled the next month as rotation groups 2, 3, and 4; and households in rotation groups 5, 6, and 7 will be sampled as rotation groups 6, 7, and 8. Households who were initially in rotation group 4 rotate off for eight months, and households in rotation group 8 have finished being in the survey. This means that—in theory—three-fourths of sample individuals can be matched from one month to the next.

In practice, however, the fraction is somewhat less than three-fourths, for a few reasons. First, households that move locations are not tracked to their new residence. Rather, the new household that occupies the housing unit (if any) is entered into the sample, with an appropriate flag in the data. Second, even if a household matches, it is not necessarily the case that everyone in the household will, as household composition can be fluid, with members entering or leaving. Third, it is sometimes not possible to tell that an individual is the same between adjacent months because some personal characteristic, such as sex, age, or race/ethnicity, changes unexpectedly; this could be due to a data mistake, but it could also result from a new person being surveyed. In these last cases, it's not clear whether these individuals should be matched.[4]

Madrian and Lefgren (2000) recommend matching individuals in the same household on sex, race/ethnicity, and age. Following such an approach and using more-recent data, Drew, Flood, and Warren (2014) show that the actual match rate among individuals from one month to the next is about 71 percent rather than 75 percent; this represents about 95 percent (71 / 75) of the ideal case. Of the missing 5 percent, about 4 percentage points are due to a household or an individual not being present the following month, and the remaining 1 percentage point is due to a faulty match on the demographic characteristics.

Nonetheless, the vast majority of individuals in the CPS can be matched from one month to the next using the approach found in the papers referenced above.

A.2 Identifying New Hires

The CPS has historically asked dozens of questions about employment and unemployment, covering a wide variety of topics ranging from the number of hours worked per week to the job search activities among the unemployed. Responses to several of these questions are used to create a new variable capturing the individual's overall employment status, distinguishing employed, unemployed, and not in the labor force.[5] Beginning with a redesign of the survey in 1994, three new questions were added in rotation groups 2–4 and 6–8 that asked individuals who reported being employed in the previous month as well as the current month whether they still worked for the same employer, whether their job activities and duties were the same, and whether the occupation and work activities reported last month were still accurate for the current month.[6]

The information in these questions allows one to identify individuals who take a new job, but it is important to be clear about what exactly should constitute a newly hired worker. It seems obvious that someone who reports not being employed in the past month and employed in the current month—that is, transitioning from nonemployment to employment—should be counted, and the longitudinal matching of the CPS makes this straightforward using the overall

---

[4] Another reason is that, even if there would be a match, a household may refuse to participate in the next month's survey.

[5] This variable, *pemlr*, is consistently coded for the entire period over which the NHQI is calculated and actually has seven values; for concision, they have been grouped into three.

[6] These variables are *peiodp1*, *peiodp2*, and *peiodp3*. Fallick and Flesichman (2004) were among the first to use these variables to measure labor flow dynamics in the CPS.

employment status variable. If someone reports being out of the labor force or unemployed in the initial month but employed in the next month, they are counted as a new hire.

Another group of newly hired workers are those who are employed from one month to the next but who switch employers. Indeed, this type of new hire is half as frequent as a transition between nonemployment to employment. These hires can be determined by individuals who report that their employers are *not* the same as in the previous month.

A trickier case consists of individuals who report steady employment and the same employer in both months, but also report that their job activities and duties have changed or that their occupations are different. These individuals may have formally applied and been hired for a new position for the same employer, received a promotion (or demotion), had their work duties changed because of a reorganization, or merely assumed different responsibilities. Unfortunately, the data make it almost impossible to tell which of these occurred. It is not clear that a change in responsibilities or duties constitutes a new job, let alone a newly hired worker. A promotion or demotion likely is closer to implying a new job, but not necessarily a new hire. Arguably, taking a job within the same company after an application process should count as a new hire (what is sometimes called an internal labor market), but there is no way to gauge how formal or informal this process is. Because of the difficulty in accurately classifying these scenarios, a conservative approach is taken of *not* counting any of these individuals as new hires. While this possibly leads to missing a few people who should be counted, it also avoids including a larger number of people who should not be counted.

Thus, for constructing the NHQI, a new hire is an individual who from one month to the next either (a) transitions from nonemployment to employment or (b) is employed steadily but reports a change in employers. In both cases, classification requires matching the same individual in adjacent months as specified above.

Of course, to be counted as a newly hired worker, it makes sense that an individual has to be *hired by someone else*, so workers who report being self-employed in the current month are not counted. Additionally, the small number of people who report working without pay, mostly individuals working in a family business but also those taking unpaid internships, are excluded, as these jobs are often temporary by construction, and the purpose of the NHQI is to capture the earnings potential of new jobs.[7]

A.3 Caveats and Correlations

The approach just described will capture new hires who transition from nonemployment to employment or switch employers *if* they stay in the same household. Recall that the CPS cannot match individuals who move or otherwise change households, and thus the approach will necessarily miss people who move and immediately start a new job. It turns out that about 1 in 10 household moves is because of a new job or job transfer, and slightly less than 1 in 8 Americans moved houses over the past year (Ihrke 2014). This means that about 1 in 80

[7] Approximately 0.4 percent of new job takers (excluding self-employment) are for jobs without pay.

Americans moved in the past year to take a new job (or be transferred), or about 2.1–2.2 million adults. Some of these individuals will get counted as new hires if the job does not start immediately after the move (and they were initially surveyed in a month before the job started). But many of them will be missed. How big a share of all new hires is due to these movers?

If the number of new hires that can be identified in the CPS, as defined above, are added up over a year, applying the sample weights, the total number of new hires is about 61.8 million.[8] If all the potential new job movers are added, the total is 64 million. Thus, at most, the movers can account for 2.2 / 64 = 3.5 percent of all new hires. The actual share is likely less than 3.5 percent, as some of the movers are job transfers (not new hires) and some are still likely to be counted. Even being unable to catch newly hired workers who move, the matched CPS approach will capture over 96 percent of new hires.[9]

The CPS, however, is not the only source that can provide a monthly estimate of the quantity of new hires. The BLS has since the end of 2000 conducted a monthly survey of job flow dynamics, called the Job Openings and Labor Turnover Survey (JOLTS). Each month, JOLTS asks some 16,000 business establishments questions about their number of job openings, hires, quits, layoffs, and other separations. Although a limited amount of information about the establishment is collected—notably its industry and location—no detail is available about the characteristics of the jobs or, for hires, the people who fill them. While JOLTS is thus not a great choice to capture the earnings power of new hires, it does allow for a comparison of the volume of new hires calculated in the CPS—and used for the NHQI—to a high-quality, independent source.

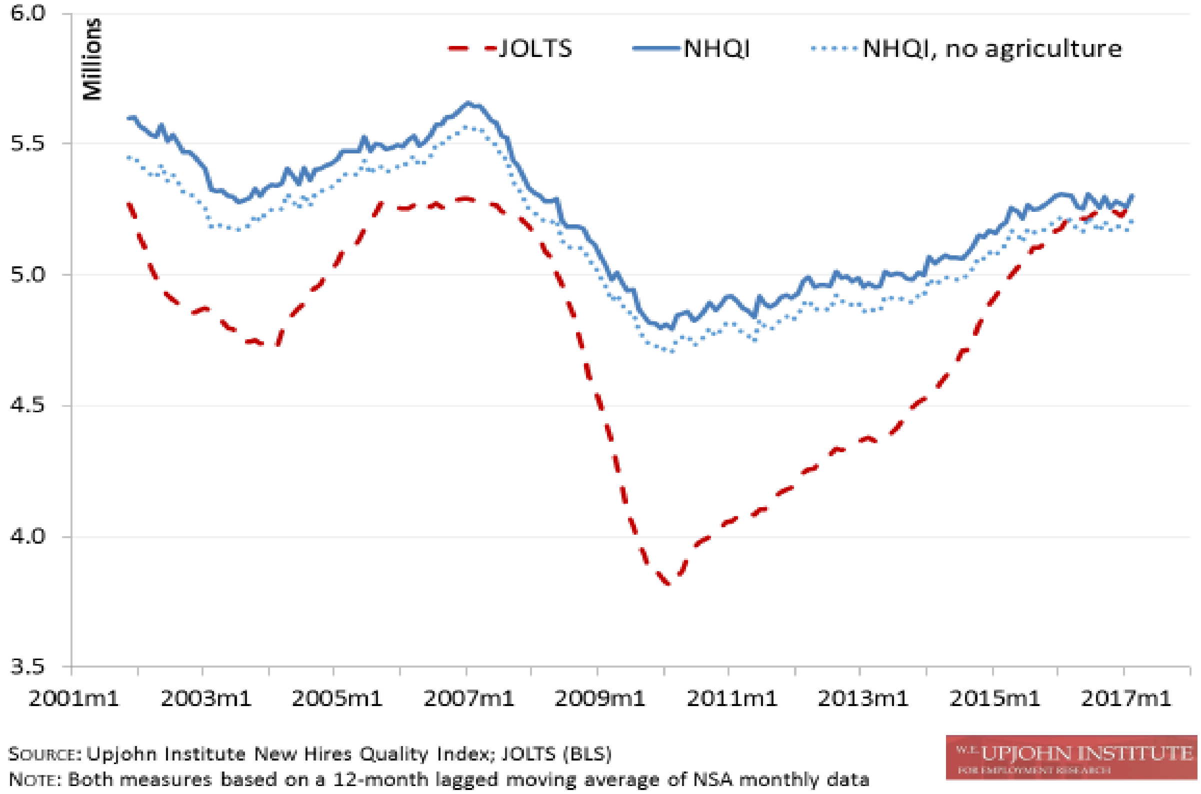
Figure 1 shows the volume of new hires each month from late 2001 through early 2017 for both the CPS-derived series (labeled NHQI) and the JOLTS series. Because the number of new hires in any particular month can be volatile, and the comparison is meant to capture longer-term differences in the data sources, each series has been smoothed by taking a moving average of the most recent 12 months.[10] The CPS-derived series, in solid blue, shows monthly volume of about 5.5 million prior to 2007, falling to below 5 million during the recession, and partially recovering to about 5.3 million by the beginning of 2017. The JOLTS series, in dashed red, tends to move similarly but generally falls a few hundred thousand below the CPS-derived series, with this gap widest during periods of weak labor demand (both the early 2000s recession and the Great Recession) and shrinking during recoveries, with essentially no gap by the end of the period. The overall correlation between the two series is quite high: $r = 0.91$.

[8] More precisely, this is the annual number of new hires as calculated from the CPS, averaged over the years 2011–2016.

[9] Since newly hired workers who move are likely to be taking higher-paying jobs than those who don't move—after all, the job must justify the cost of the move—the approach will disproportionately miss some of these higher-paying jobs. As long as this phenomenon is stable over time, however, it will not affect *changes* in the NHQI.

[10] Since a 12-month moving average obviates the need to seasonally adjust the data—hiring is routinely more common at certain times of year than others—this approach is used throughout.

Figure 1 Comparison of New Hires Volume: NHQI and JOLTS



The relative closeness of the two measures is quite reassuring, especially since there are conceptual reasons for why the CPS-derived series should be both slightly higher and less sensitive to economic forces. First, the sampling universe of JOLTS excludes agricultural jobs while the CPS universe includes them (excepting family farm jobs). This difference can be seen in the dotted blue line, which plots the CPS-derived series without these agricultural jobs, and closes the CPS-JOLTS gap slightly. Second, the CPS is more likely than JOLTS to capture certain very short jobs that last less than one month; this results from the different reference timing in the two surveys. Third, the household-based CPS is also more likely to capture unofficial or casual forms of employment than the business-based JOLTS.[11] Because such informal jobs—ranging from day labor to babysitting to app-based gig work—become more common during downturns relative to formal employment (Bernhardt and Thomason 2017), it's not surprising that the gap between the CPS and JOLTS grows during those periods. All in all, the volume of new hires picked up in the CPS seems quite reasonable relative to other measures of hiring.

A.4 Occupations

Even once individuals have been matched and new hires identified, construction of the NHQI requires knowledge of their occupations. The CPS has long asked for respondents' occupations, but as new types of jobs have evolved, the coding of occupations has also had to change in the

[11] JOLTS also may miss businesses that are highly seasonal or short-lived (Davis et al. 2010), while the CPS would ostensibly capture employment at these places.

survey.[12] These changes typically happen around decennial censuses, as the CPS uses a nearly identical coding as used by the long-form census, and its successor, the American Community Survey. Since the beginning of 2011, the occupational coding in the CPS has used 2010 census codes. Between 2003 and 2010, 2000 census codes were used; and prior to 2003, 1990 census codes were used. Some of these coding changes were more significant than others; in particular, the switch from the 1990 coding to the 2000 coding was substantial, as many new computer-related jobs were recognized, and attempts were made to harmonize the U.S. coding system with that of other countries for easier international comparisons.

A consistently defined set of occupations is important, because the NHQI requires merging in wage data from another source (described below) based on the modern classification of occupations. In fact, the modern system, the Standard Occupational Classification (SOC), is the evolving result of trying to harmonize U.S. occupational codes with international systems and nest the existing census codes. That is, the SOC provides greater granularity than the census codes used in the CPS, but they can hierarchically nest into them.[13]

For the period 2011–2017, it is straightforward to convert the occupation codes in the CPS to the modern (2010) SOC system, and the IPUMS organization at the University of Minnesota Population Center provides a crosswalk. For the period 2003–2010, the CPS occupation codes must be crosswalked *twice*: once from the 2000 census coding scheme to the 2000 version of the SOC system, and then again from the 2000 SOC to the 2010 SOC. For the first bridge, IPUMS again provides a crosswalk. However, there is no readily available crosswalk for the second bridge, between the 2000 and 2010 SOC systems.[14] This latter change was relatively minor, with several occupations assigned new codes, and a few occupations that were combined; both are simple to map. Somewhat trickier is that a few other occupations were split into new (closely related) occupations. These splits are randomly assigned based on the empirically observed shares in the 2010–2012 ACS.[15] While this random assignment is not perfect, because the splits constitute a small share of employment or new hires, and related occupations are typically similarly paid, mistakes are likely to have a minimal effect on earnings power.

This crosswalking strategy yields consistent occupational coding in the CPS from 2003 to date, but recall that new hires can be identified in the CPS from 1994 on. It would be desirable to include data from this period as well. Unfortunately, it is exceedingly difficult to convert 1990

[12] In practice, the survey asks respondents to describe what they do, and a paid coder translates that description into one of several hundred designated occupations.
[13] For example, there are 820 defined occupations in the 2010 version of the SOC, but 532 defined occupations in the 2010 census coding. In several cases, multiple (more-detailed) SOC occupations map into a single census occupation, but an SOC occupation will not be split across multiple census occupations.
[14] Technically, BLS does provide a crosswalk, but it does not account for the loss of granularity that occurs when converting the 2000 census codes to the 2000 SOC system. After this mapping, there is no longer a 1:1 relationship between the codes in the two SOC systems.
[15] By comparing large, multiyear samples of the ACS immediately before and after the coding change, it is possible to observe how much of an occupation splits into each of its successor occupations. These shares are used to probabilistically assign a 2000 SOC code in the CPS to a 2010 SOC code. Of the several hundred SOC occupations, only a dozen or so are split, and their total share of new hires is quite small.

census occupational codes to 2000 census occupational codes, and there are no 1990 SOC codes, as that system did not yet exist. The census does provide a crosswalk between the 1990 and 2000 census codes (Scopp 2003), but almost every occupation in one system is split into multiple occupations in the other (a many-to-many mapping). Such a probabilistic crosswalk would create unacceptably high amounts of misclassification.[16]

Fortunately, there is a partial solution. The Census Bureau and BLS released files for the CPS from 2000 through 2002 that included codes for occupation, industry, and a few other variables that were updated to use the 2000 census standard.[17] These revised files (accessible through the National Bureau of Economic Research) can be merged to the original CPS files from 2000 through 2002, and the above crosswalking procedure for the 2000 systems can be used to create a consistent occupational coding from 2000 to the present. It is not possible, however, to convert occupation codes from the 1994–1999 CPS to the modern SOC system with a high degree of reliability, and thus the NHQI can extend back only to 2000.

A.5 Industries

While occupations capture a worker's tasks and provide highly meaningful information about earnings, it should be apparent that not all workers in a particular occupation have the same skills or even do precisely the same things in the same environments. For example, broadly speaking, an electrician works to install, maintain, and repair electrical equipment and wiring. Many electricians work in construction of new homes and need the knowledge and skills to make sure that wiring is up to residential code. Other electricians may specialize in ensuring that telecommunications equipment is working properly and need to understand the operating specs for sophisticated broadcast transmitters. These differences in skills and activities on the job can translate into differences in typical pay. Whereas the typical electrician working in construction earns about $24.50 per hour, the typical electrician working in the media and publishing sector earns about $36.40 per hour.[18]

To capture—at least partially—these differences in earnings power even within occupations, one can also exploit the industry information available for workers in the CPS. Like occupations, industry codes also change over time and need to be harmonized. Unlike occupations, however, industry harmonization at the finest level of detail is not pursued; rather, harmonization at the

[16] IPUMS also provides a crosswalk from the 1990 census codes to 2010 census codes. However, rather than being probabilistic, this simpler crosswalk assigns to an older code the modern code of the largest-share occupation that it maps into. For example, the very large 1990 census occupation "Managers and administrators, not elsewhere classified" is split into over 30 2000 census occupations, the largest of which is "Managers, all other," constituting just 28 percent of the 1990 code. Since the other occupations range from the very highly paid (chief executives) to the modestly paid (secretaries and administrative assistants), there is large scope for error. In fact, the 499 occupations in the 1990 census system map to only 352 out of the roughly 500 occupations in the 2010 system, leaving about one-third of available occupations unassigned.
[17] Weights were also revised to reflect the 2000 census.
[18] These figures come from the OES data source, to be discussed shortly. The differences in pay may also reflect other factors besides skill specialization, such as geographic location, employer size, and unionization rates.

sector level, of which there are about 20 categories, is adopted.[19] This choice reflects both feasibility and practicality. First, modifications to detailed industries occur more frequently than they do to detailed occupations; the first happens roughly every five years, while the second happens roughly every 10 years. Second, many detailed occupations may have no workers at all in a detailed industry (e.g., there are no surgeons working in fast food restaurants). Third, earnings differences within an occupation for closely related industries are likely quite small (e.g., electricians working in house construction use very similar skills to those working in commercial building construction).

At this level of aggregation, harmonization of industries over time is straightforward, using the detailed industry codes provided in the CPS, especially with the supplemental files for the years 2000 through 2002.[20]

B. Occupational Employment Statistics

The Occupational Employment Statistics program run by the Bureau of Labor Statistics is the primary (and only source) of employment and wage information on occupations that is asked of businesses. Twice a year, the Bureau, in conjunction with state workforce agencies, survey some 200,000 business establishments across the United States about the occupations of their employees, using the same SOC coding system described above. In addition to the count of employees by occupation at each establishment, the survey asks the number that fall into each of 12 wage bins. By aggregating six waves of the survey over three years (and adjusting wages for inflation), the OES program is able to estimate employment count and wage distributions for over 800 occupations not just for the country as a whole, but for states and metro areas as well as breakdowns by industry.[21] These data are released annually, as two new waves replace the oldest two waves in the three-year moving average. The NHQI uses the latest release, for 2016, which was issued in May of 2017.[22]

For each occupation (and area or industry), the OES data provide information on the mean or average wage, as well as the 10th, 25th, 50th (median), 75th, and 90th percentiles, both at hourly and annual rates.[23] These wage data can be merged by detailed occupation code to new hires in the CPS, but the process requires a few decisions.

---

[19] This level corresponds to the two-digit North American Industry Classification System (NAICS) and is roughly the grouping used for industries in monthly employment reports from BLS; examples include manufacturing, retail trade, and finance and insurance.
[20] This harmonization crosswalk for industries is available upon request.
[21] For more information, see https://www.bls.gov/oes/.
[22] Note that using a single release of OES wage data means that occupational wages will not vary over time or reflect differential wage changes across occupations. This issue will be addressed in later updates to the NHQI.
[23] A few occupations that generally don't have year-round, full-time work may report only hourly or only annual wages. In the latter case, hourly wages are constructed by assuming year-round full-time employment (dividing annual wages by 2,080 hours.)

First, what wage statistic should be used? Since the survey covers incumbent workers, including many with extensive experience or tenure, the mean or median may not be appropriate to capture the earnings power of newly hired workers. On the other hand, some new hires are relatively senior, and the 10th percentile may be too low. Thus, the 25th percentile—the wage where one-quarter of workers earn less and three-quarters earn more—was chosen as the most appropriate wage measure.

Second, what breakdown by area or industry, if any, should be used? The national occupational wage data is more reliable due to larger sample sizes, but it overlooks some substantive differences in wage within the same occupation across industries, as mentioned above. Breakdowns by state or metro area would also capture within-occupation wage variation, but may be less reliable, especially for small areas, where sample sizes shrink. As a compromise, occupations broken down by industry-sector (two-digit NAICS) were chosen.[24]

Finally, whereas over 800 occupations are represented in the OES, only about 500 occupations are coded in the CPS; the latter are coarser versions of the former, in which some relatively small occupations are combined. Because the SOC system is hierarchical, this type of nesting is straightforward. To preserve maximum detail, the occupational merging process follows the hierarchy, with an initial merge at the most detailed level of occupation (six-digit SOC). This first pass matches most occupations but leaves several of them in the CPS unmatched because of coarser occupations. The next pass attempts to match at a slightly higher level, five-digit SOC codes.[25] This proceeds to higher levels of aggregation until all the occupations of new hires in the CPS have matched wage data from the OES.

C. Demographic Adjustment

This matching process will assign wages to all new hires, but the only source of variation will be the differences in occupations. Put differently, the wage assigned to a 20-year-old licensed practical nurse on her first job will be the same as one assigned to a 35-year-old veteran LPN switching hospitals. Clearly, it would be desirable to adjust for such demographic differences among new hires. Moreover, since variation in earnings power due to experience, education, and other characteristics can be different across occupations, adjustments should incorporate these occupational differences.

The detail in the CPS allows for these types of adjustments using a multistep regression-based approach. The idea of the regression is to relate the self-reported wages among CPS respondents to both their own characteristics and the context of the survey. Of course, since the CPS contains self-reported wages, a natural question is, Why go through all the effort of matching on occupations when actual wage data are available? This question will be answered

[24] In practice, the choice of more detailed occupation breakdowns affects the *level* (hourly wage) but not the trend of the index. That is, once normalized to a common starting point, an index based on overall occupation data and one based on industry or area breakdowns look very similar.

[25] The OES provides wage data at various levels of occupational aggregation, from two-digit occupation classes (e.g., construction and extraction workers) to full six-digit detail (e.g., pile-driver operators).

thoroughly in the comparison section later in the report, but the short answer here is that valid self-reported wages exist for only a small fraction of new hires, and while this fraction is generally too small to make for a reliable index on self-reported wages alone, it is sufficient when pooled across several years for use as a dependent variable in a regression.

The adjustment procedure occurs in three steps. In the first step, the natural logarithm of inflation-adjusted hourly wages of individuals are regressed on nondemographic variables from the survey, pooling across the entire sample horizon from 2000 through the latest month of data and using sample weights.[26,27] In the second step, the residuals from this regression become the outcomes for a series of additional regressions on individual demographics. In particular, the residuals are regressed—*separately* for each of 110 occupation groups (four-digit SOC)—on indicators for sex, race/ethnicity, education, and a quartic in age. These separate regressions allow the relationship between demographics and wages of new hires to vary by occupation, with some occupational aggregation necessary to balance sample size. In the third step, the predicted values are exponentiated (to undo the log transformation) and multiplied by the occupational wages taken from the OES.

This process allows differentiation in predicted wages not only between the two LPNs given in the example above, but also for many other demographic differences, such as a retail store manager with an associate's degree and one with a bachelor's, or other aspects of pay differences by sex and race.[28] The role that this demographic adjustment plays will be explored further in the section on robustness, below.

D. Final Processing

After the construction of these adjusted wages for all new hires, the mean is calculated for each month, both overall and for a variety of subgroups taken from the individual characteristics in the CPS. These subgroups include breakdowns by sex, age groups, education, public and private sectors, goods and service-producing industries, region of the country, and type of new hire (from nonemployment and changes in employer).

---

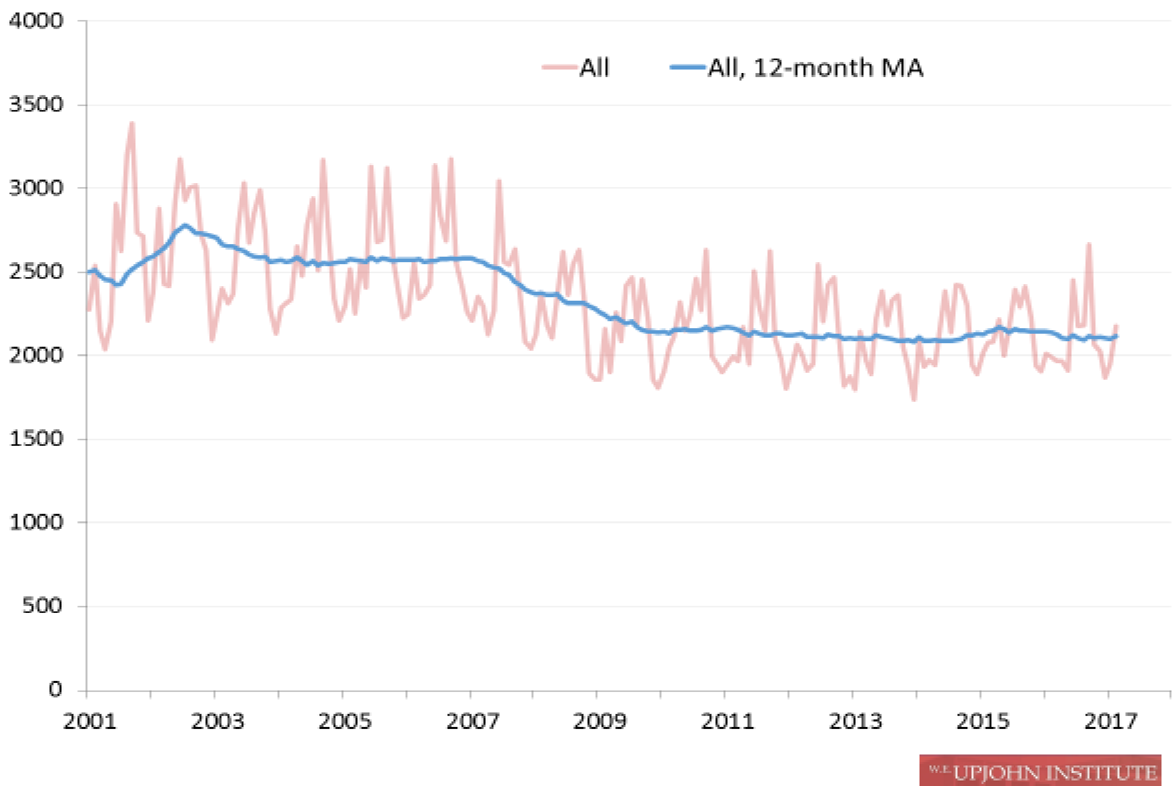
[26] The construction of this wage measure follows standard practice in the academic economics literature (see, for example, Acemoglu and Autor [2011] for a review). Among workers who identify as hourly, hourly wages are directly self-reported; for salaried workers, the weekly salary is divided by the usual hours worked per week. In both cases, wages that are imputed by the CPS, rather than directly self-reported, are dropped. The hourly wages are adjusted for inflation to year 2015 dollars using the Personal Consumption Expenditures deflator from the U.S. Bureau of Economic Analysis, and then logged. The regression excludes inflation-adjusted hourly wages less than $2 or greater than $100.

[27] The control variables include indicators for all observed months, month in sample, class of employee (government, for-profit, nonprofit), type of new hire (from nonemployment or change in employer), detailed (six-digit SOC) occupation, and three-digit industry.

[28] These latter differences could result from discrimination, but they could also stem from factors not controlled for, such as firm size and location or the value of nonwage amenities. Whatever the reason, the adjustment process helps capture observed differences.

However, even these means by month can be volatile because of relatively small sample sizes in the number of hires (especially for subgroups) and to seasonal patterns, such as summer and holiday hiring spikes. Thus, it is desirable to smooth these fluctuations to understand the underlying trend. One such smoother that is straightforward, intuitive, and easy to implement is a 12-month lagged moving average.[29]

Figure 2 NHQI Sample Size over Time, Raw and Smoothed



SOURCE: Upjohn Institute New Hires Quality Index; CPS

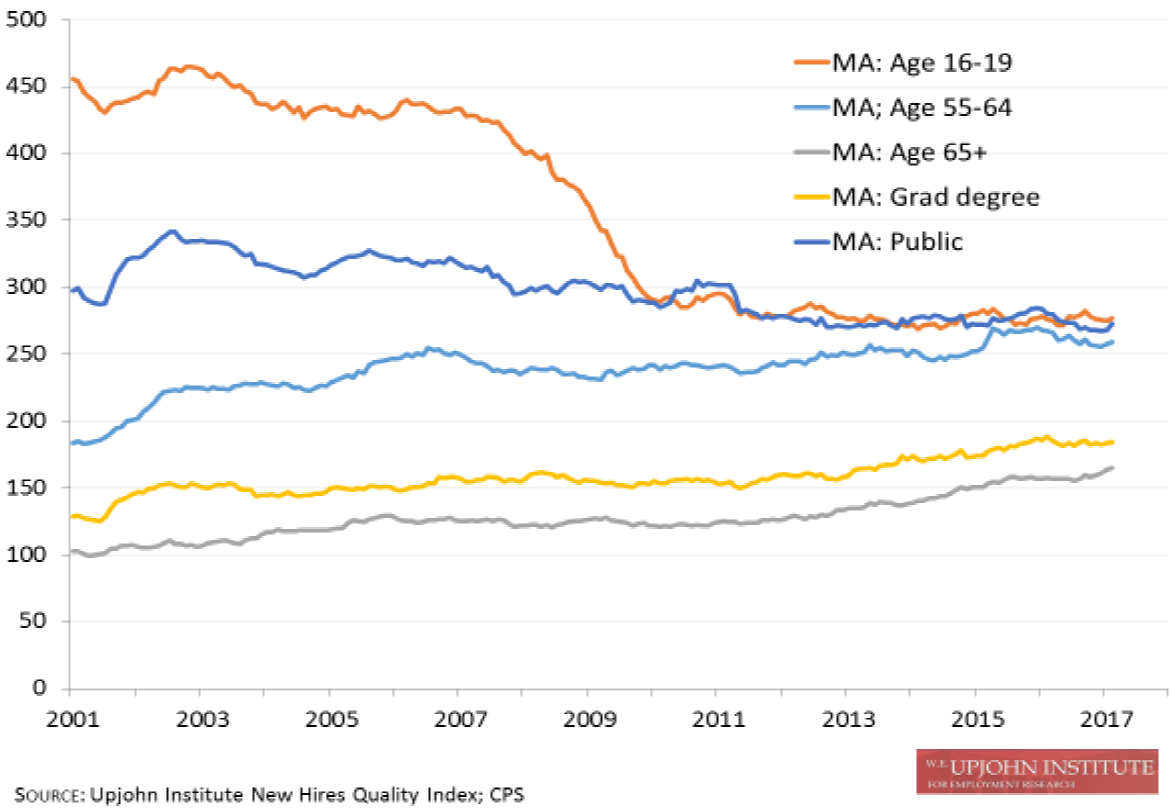
Figure 2 shows both the raw sample size of newly hired workers in the CPS and a 12-month moving average smoothed size from 2001 through the beginning of 2017. The salmon-colored line shows both the volatility and seasonality in the raw counts, with the spikes occurring in the summer and early fall months, and the troughs in the middle of winter. Prior to the Great Recession there were slightly over 2,500 individuals in the CPS who were newly hired on average each month, although the number bounced between 2,000 and 3,000. From the Great

[29] Other approaches were considered, including the official seasonal adjustment process used by the U.S. Census Bureau, the X-13 ARIMA-SEATS process, and a Hodrick-Prescott filter. Both these other methods are complicated and less transparent. Although the X-13 process reduced seasonality, significant volatility from month to month remained, perhaps because much of the volatility is due to factors besides seasonality. The Hodrick-Prescott filter reduced volatility and has the advantage of capturing more quickly changes in the underlying trend of the data and producing estimates for the full period rather than losing them to the needed lag structure of the moving average; however, in practice it produced very similar results to the moving average. Thus, the moving average was adopted for simplicity.

Recession onward, there was a decline in this count of about 500 per month. Although these raw counts do not represent the population of hires, they accord with the well-documented decline in hiring rates even as the economy has recovered (Davis and Haltiwanger 2014; Hyatt and Spletzer 2013).

As noted earlier, an advantage of the CPS is that hires can be broken out by demographic group. With an overall sample size of roughly 2,300 new hires per month, there is sufficient power for several different breakdowns. Figure 3 shows smoothed monthly sample counts for a selection of different groups.

Figure 3 NHQI Smoothed Sample Size over Time, by Select Group



SOURCE: Upjohn Institute New Hires Quality Index; CPS

Teenagers initially have the largest counts, reflecting their rapid turnover in jobs and relatively short tenures. During the Great Recession, the raw hire counts (and weighted ones, too, although not shown) fell substantially and have not recovered. In contrast, raw new hire counts among seniors (age 65+) are initially among the lowest but rise steadily throughout the period, reflecting both greater labor force attachment and hiring dynamics among baby boomers. Note that while hiring is still relatively uncommon in this age group, the sample size is at least 100 in every month.[30] Most other demographic groups have new hires counts that fall in between.

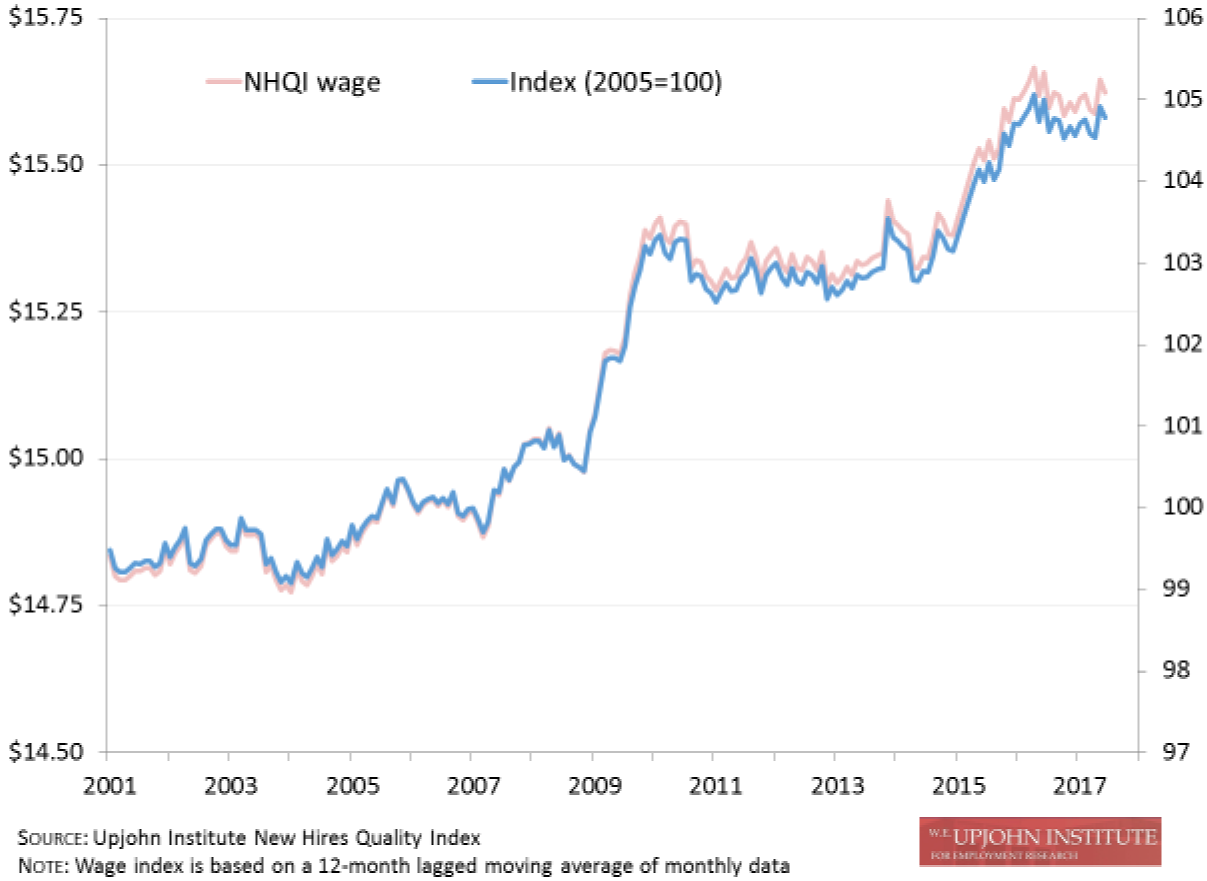
[30] When applying a 12-month moving average, the effective sample size used to create the NHQI is 12 times as large, so even for seniors, more than 1,200 observations will contribute to the index, which yields greater reliability.

III. The NHQI

A. Overall NHQI

The methodology above yields an hourly wage measure based on a 12-month lagged moving average of demographically adjusted occupational wages of new hires: the NHQI. This index can be expressed in terms of hourly wages, or, to make changes over time easier to interpret, it can be normalized to a base period. Figure 4 plots both versions, with the salmon line tracking the hourly wage measure (left axis) and the blue line tracking the normalization that sets the average value in the year 2005 to 100 (right axis). Thus, the normalization shows the percentage change in the NHQI since the base period, 2005. In the figure, the axes have been scaled so that the two series nearly overlap each other.

Figure 4 NHQI: Hourly Wages and Normalization (2005=100)



As of June 2017, the NHQI hourly wage was $15.62, an increase of $0.79 from $14.84 in the first month of the NHQI, January 2001. Although mentioned earlier, it is worth repeating that these figures do NOT represent the average actual wage of all new hires in those months. Rather, the NHQI captures changes in earnings power due to the occupations into which people are being hired and changes in the demographics of those hires. For this reason, *changes* in the NHQI are more useful than *levels*. Looking at the normalized blue line, the NHQI as of June 2017 is 104.8;

that is, the index is 4.8 percent above its value in 2005. (It is 5.3 percent above its value in January 2001.)

Examining the NHQI more broadly, there was little growth between 2001 and 2005, mild growth between 2005 and 2009, and then very sharp growth in the tail end of the recession and its immediate aftermath. During the recovery, growth was essentially flat until 2015, when the NHQI again grew quickly for about one year. Over the past year and a half, the index has again been relatively flat.

It may seem surprising that the NHQI grew so rapidly during the end of the recession, as employment continued to fall. Shouldn't wages rise in a boom, not a recession? It is important to remember that the NHQI is constructed from newly hired workers, and Figures 1 and 2 suggest that the volume of these workers fell during the recession. In fact, Figure 5 shows this contraction in hiring volume more directly by plotting the aggregate hiring volume (and normalized to 2005) of the NHQI.

Figure 5 NHQI: Monthly Hiring Volume and Normalization (2005=100)

Over the Great Recession, hiring volume fell by 15 percent. While it has recovered somewhat, it is still below its previous peak and its level in 2005, and in fact is roughly at the same volume reached in the trough from the early 2000s recession. That the NHQI wage rose sharply over the recession as hiring volume fell implies that the composition of individuals who managed to get hired during the recession was quite different than those who got hired before or after.

More succinctly, individuals hired during the Great Recession were more highly skilled, taking higher-paying occupations and having greater demographic advantages (age, education) than previously.

Perhaps a larger surprise is that during the recovery as hiring volume began to grow, the NHQI wage did not fall but instead plateaued. If hiring into low-skilled occupations or among those demographically disadvantaged was muted during the recession, and this reversed during the recovery, one would expect the NHQI wage to fall accordingly; that it did not suggests that employers may permanently be seeking higher-skilled workers. Indeed, recent research finds that employers increased their skill demand in job postings during the recession, and that this effect persisted throughout the recovery, which is consistent with an episodic restructuring in the demand for labor (Hershbein and Kahn 2017).[31] It also helps explain the so-called jobless recovery, in which many workers who were laid off could not find reemployment, even as the economy strengthened.

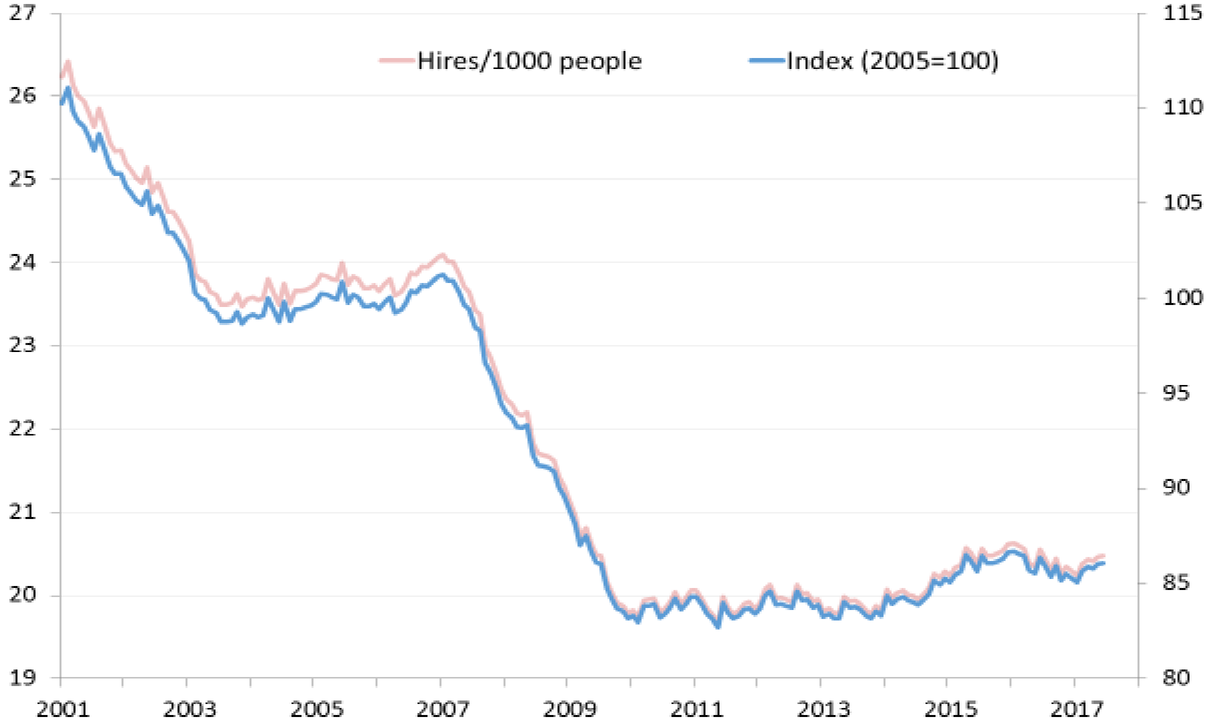
Another facet of the quantity of jobs, however, is observed by benchmarking them not to earlier aggregate magnitudes, but to the (changing) population. That is, rather than present hiring volume in millions of hires per month, it is possible to show hiring rates—the number of newly hired workers per thousand people—thus accounting for population growth.[32] Figure 6 shows these hiring rates. From an initial high of about 26 hires per thousand people in 2001, the hiring rate fell to about 23.5 hires per thousand in the aftermath of the early 2000s recession; it barely recovered to 24 hires per thousand in 2007 before plunging below 20 hires per thousand in 2010, and has barely gained ground since. Although part of this decline is due to an aging workforce, the stair-step pattern suggests that recent recessions are having long-lasting effects on hiring dynamics (Hyatt and Spletzer 2013).

It can also be useful to combine both the NHQI and hiring volume (in levels, not rates) into a summary measure by taking their product. The result is the wage bill, or the cumulative hourly earnings power for all workers hired in a given month. This series is shown in Figure 7. Although this figure looks similar to the volume shown in Figure 5—changes in hiring volume vary more over time than changes in the wage index—there is an important difference.

[31] Later, it will be shown that the plateau in the NHQI during the recovery is not due to demographic adjustment but the actual occupations into which people are hired.

[32] Note that some other data sources, such as JOLTS, consider hiring rates with the number of workers rather than the number of people as the denominator. For data reasons, we use independent population estimates from the U.S. Census Bureau rather than population estimates from the CPS, which restricts our choice of denominator.

Figure 6 NHQI: Hires per Capita and Normalization (2005=100)



SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

Figure 7 NHQI: Wage Bill and Normalization (2005=100)

SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

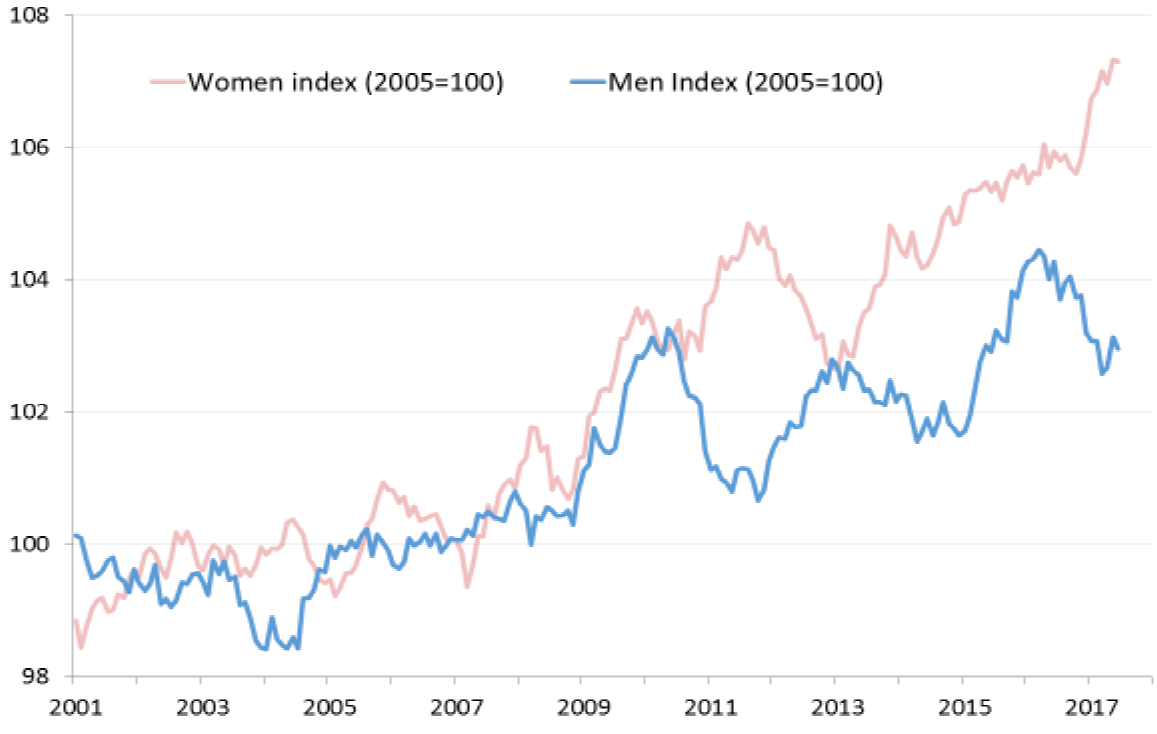
Notably, while volume has barely recovered to the low reached during the early 2000s recession, the wage bill has easily surpassed this trough and its level in 2005, and is actually close to fully recovering to its pre–Great Recession high. Indeed, the NHQI wage bill measure provides a powerful summary measure of overall hiring activity, encompassing both the quality and quantity of new jobs taken.

To summarize, there are four distinct elements of the NHQI. The main element is the hourly wage index shown in Figure 4. The three ancillary elements help give the hourly wage index additional context: the hiring volume, the hire rate, and the wage bill.

B. Subgroup NHQI

One of the strengths of the NHQI is that the different elements can be constructed not just for all workers but for a multitude of demographic groups. Thus, the NHQI hourly wage index for men can be compared to that of women, or the wage bill of bachelor's degree holders can be compared to the wage bill of workers with other levels of education.

Figure 8 NHQI: Wage Index for Men and Women (Normalization 2005=100)



SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

W.E. UPJOHN INSTITUTE

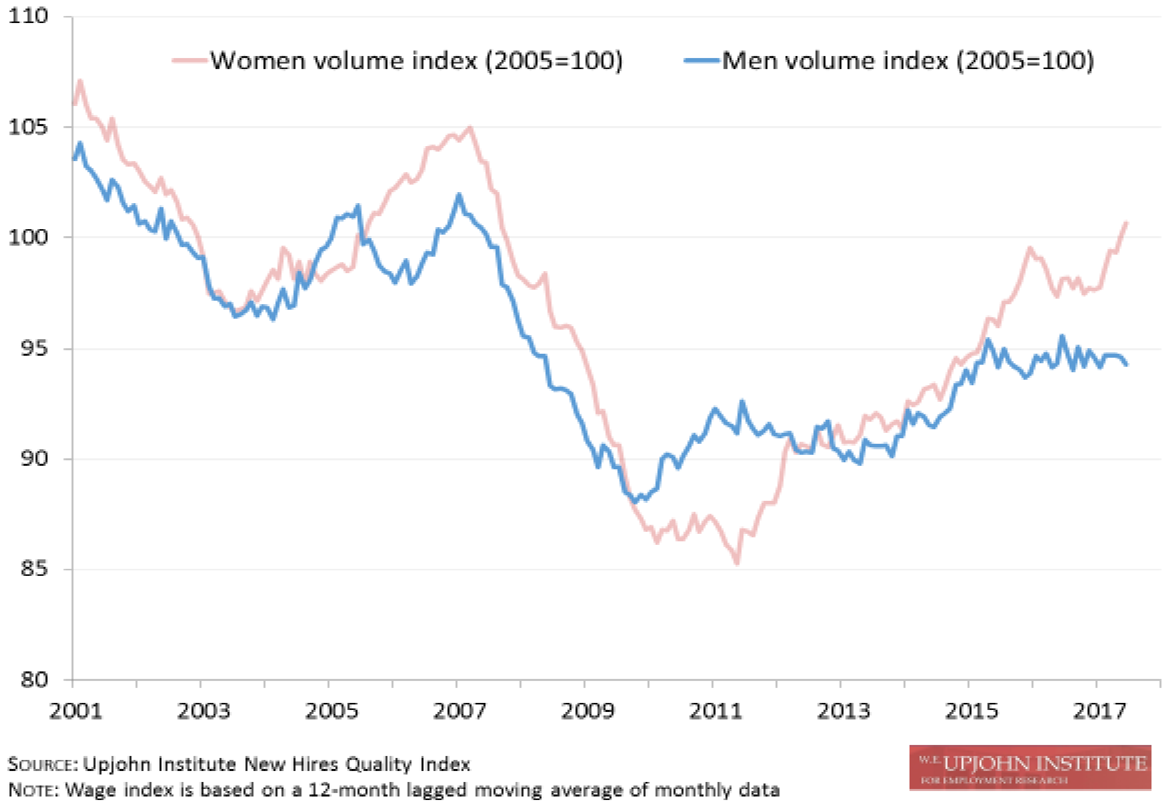
Figure 8 considers the first example of men and women. To ease comparison, the series are normalized so that the values equal 100 in the year 2005 for both sexes.[33] The most dramatic

[33] In levels, the hourly wage index for men is higher than that for women throughout the period.

pattern from the figure is the stronger growth for women compared to men. The NHQI hourly wage index for women is 7.3 percent above its 2005 level, whereas for men the index is only 3.0 percent above its 2005 level. Most of this gap has opened since the Great Recession ended. Interestingly, besides an 18-month period in 2015 and 2016, the wage index has moved in opposite directions for men and women, likely because of changes in the occupations most in demand.

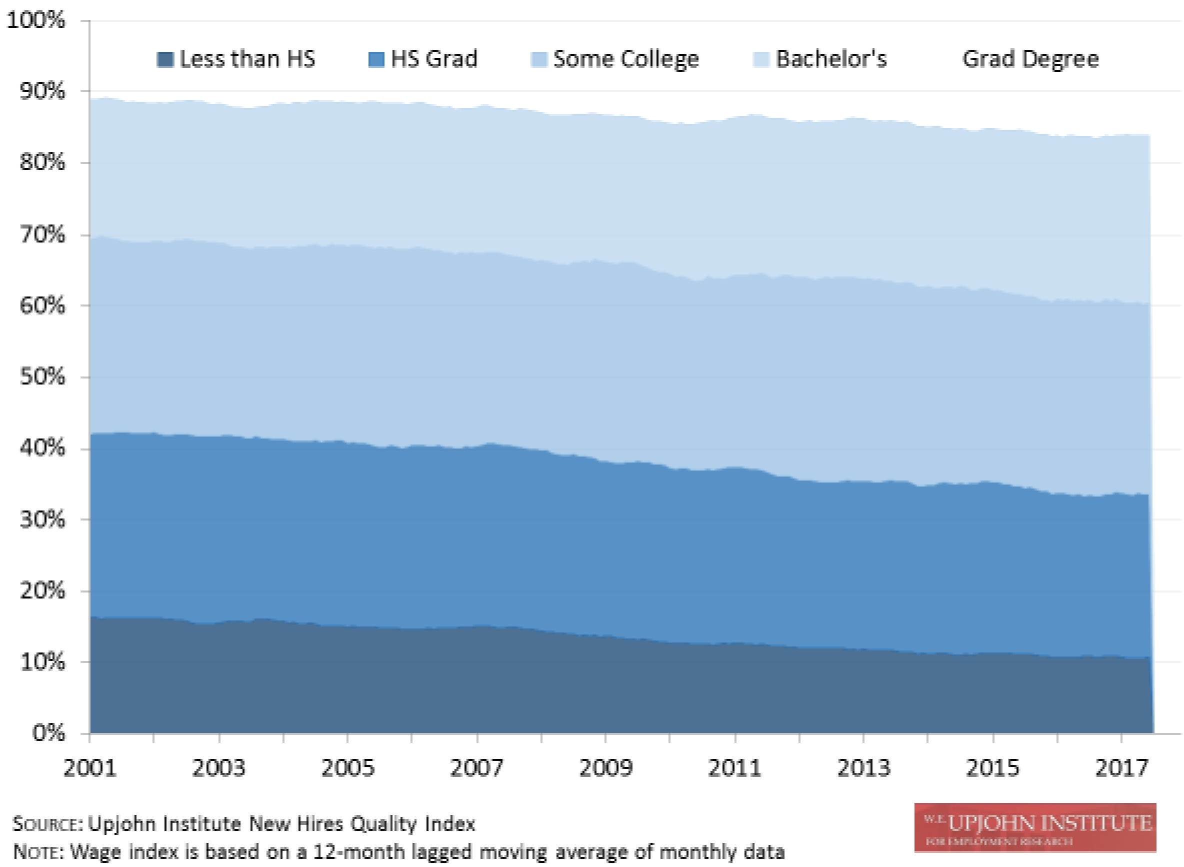
Of course, it is well known that production and construction workers were severely affected by the Great Recession. Since these workers are disproportionately men, perhaps the weaker wage index for men stems from their more severe exposure to the recession. Figure 9 demonstrates that this is not the case, at least not in terms of hiring activity. Women, not men, experienced the sharper drop in hiring over the Great Recession, and they have had a stronger recovery since. Although women's hiring volume has not recovered to its prerecession high, it is at least above its trough from the early 2000s recession; men's hiring volume, despite some recovery, hasn't even reached this low bar. Thus, relative to men, since the Great Recession women not only have been more likely to take a new job, but the earnings potential of that new job is increasing faster for them.

Figure 9 NHQI: Volume Index for Men and Women (Normalization 2005=100)



Another salient comparison is how the earnings power of newly hired workers has shifted by education. Figure 10 shows the cumulative percentage of the wage bill accounted for by different education groups. For example, the darkest-shaded area is the share of the wage bill earned by newly hired workers without a high school diploma. In 2001, this share was about 16 percent, but by the middle of 2017 it had fallen to 10 percent. The next-darkest area represents the wage bill of high school graduates, the next area that of workers with some college but less than a bachelor's degree, and the lightest area that of workers with exactly a bachelor's degree. The sum of these four areas thus represents the wage bill of newly hired workers with no more than a bachelor's degree. That this share has fallen from about 89 percent in 2001 to 84 percent in mid-2017 implies that the residual—the wage bill of workers with a graduate degree—has grown from about 11 percent to about 16 percent. (Among workers with a bachelor's degree or more, the wage bill share has increased from about 30 percent to 40 percent over this period.)[34]

Figure 10 NHQI: Wage Bill by Education



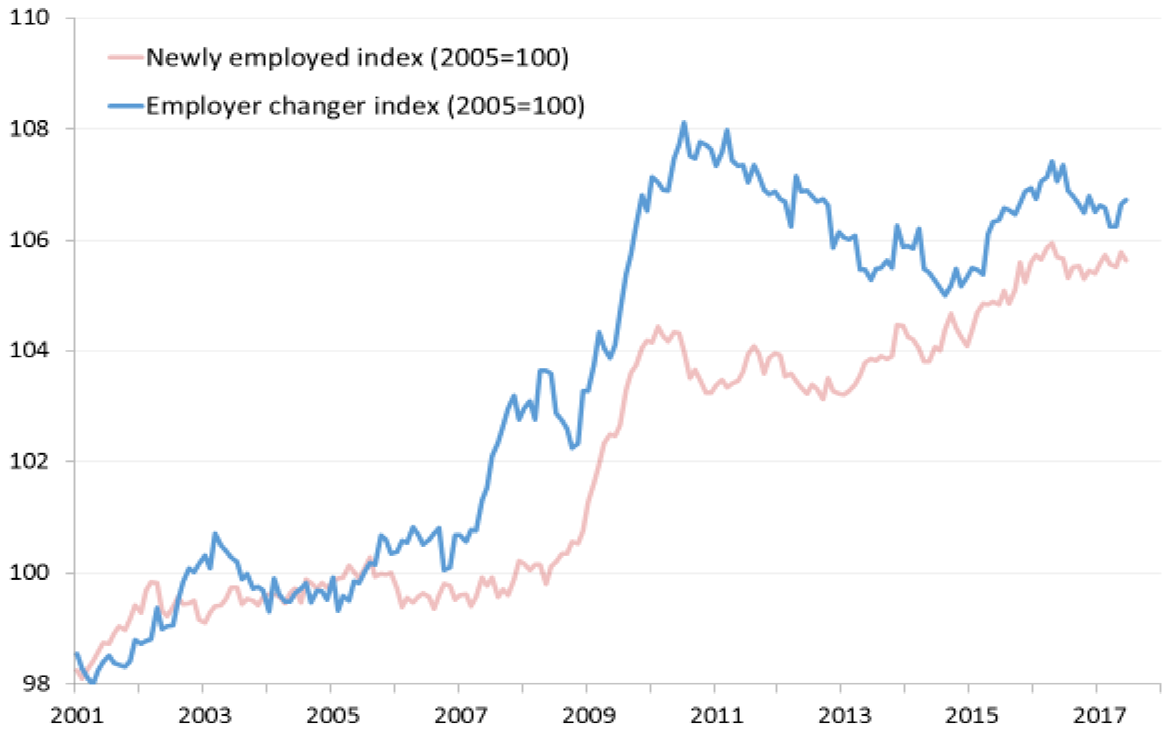
A third example of how the NHQI permits group comparisons is the case of the type of new hire. As described in part A.2 of the Methodology section, a newly hired worker is an individual who transitions from nonemployment to employment or one who transitions from one employer to another while remaining employed. These two types are quite different. The former can include people who had lost their job, spent time while unemployed looking for a

[34] These shares are in line with shares of the incumbent workforce by education, but since jobs tend to last longer among more-educated workers, this implies gains in the wage bill share among the better-educated.

new one, and (eventually) met with success. But it can also include new entrants to work (those who have not worked before) as well as reentrants to the labor force (who may have been away from work for some time while taking care of family members, recovering from an illness, or obtaining further education). The latter type disproportionately consists of workers who have quit one job for a better opportunity elsewhere, although it can also include involuntary separations (either layoffs or firings for cause) from which individuals very quickly recover. Due to the nature of these different types of changes, the former group tends to have less earnings power than the latter group, although it is a larger share of all new hires.

Figure 11 shows the NHQI wage index for each of these two hire types, with the series normalized to their 2005 values. Although the index grew faster over the Great Recession for employer changers than for the newly employed, by the beginning of 2017 the cumulative growth for the two was similar, as the newly employed caught up.

Figure 11 NHQI: Wage Index by Hire Type (Normalization 2005=100)

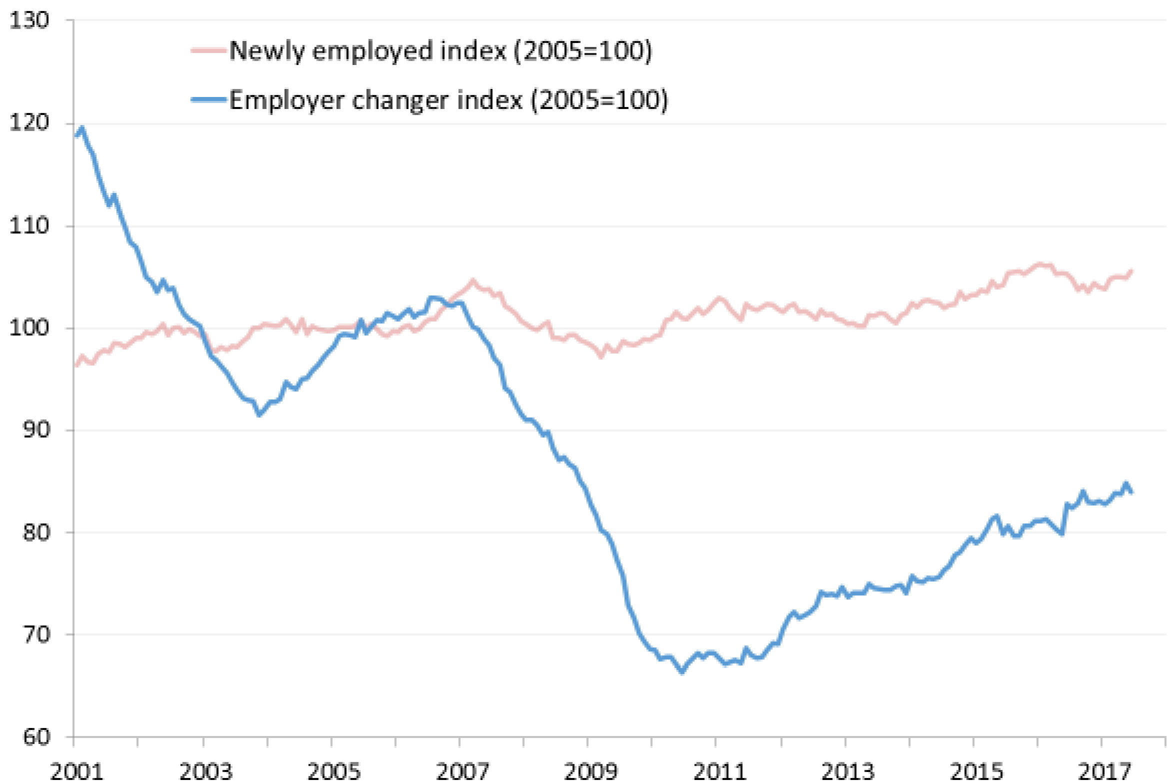


SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

To provide additional context, Figure 12 shows the (normalized) hiring volume for the two hire types. Remarkably, hiring volume among the newly employed has been remarkably stable since 2001, dipping only slightly during the Great Recession and rising modestly since to reach about 5–6 percent above its 2005 level. In contrast, the hiring volume for employer changers has seen much greater vicissitudes, falling by roughly one-quarter in the early 2000s recession, and by about one-third over the Great Recession. Despite slight recoveries following the recessions, hiring volume remained substantially below prerecession peaks, and indeed the volume of this component explains more than the entire decline in overall hiring volume shown in Figure 5. That the wage index for employer changers grew so sharply during the Great Recession (Figure 11), even as volume plummeted, is consistent with the pattern seen in Figures 4 and 5 but taken a step further: individuals who took a job during the recession by switching employers were even more positively selected than individuals who took a job at that time in general.[35]

From 2001 to the middle of 2017, the share of new hires that were newly employed rose from 58 percent to 68 percent (and the wage bill share from 53 percent to 62 percent). Nonetheless, despite the reduction in hiring activity for employer changers and the expected reduction in earnings power among new hires (Haltiwanger, Hyatt, and McEntarfer 2015), the growth in the wage index among the newly employed means that the overall earnings power of new hires has continued to grow (Figure 4).

Figure 12 NHQI: Volume Index by Hire Type (normalization 2005=100)



SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

These subgroup comparisons represent just a few of the possibilities allowed by the NHQI. The current subgroups over which the NHQI is calculated each month include sex (two groups), age group (nine groups), education (five groups), firm ownership (two groups), industry (two

[35] This selection is also evident in the slight decline in the wage index for employer changers between 2011 and 2015, which suggests that workers with lower earnings power were gradually able to begin changing employers again. Consistent with Figure 4, no such decline occurs for the newly employed.

groups), region (four groups), and hire type (two groups). Across the four NHQI elements (wage, hiring volume, wage bill, hires per capita) and subgroups, there are nearly 100 different series that constitute the NHQI, and all are available for investigation and download at www.upjohn.org/nhqi.[36]

IV. Robustness and Alternate Specifications

A. Effects of Demographic Adjustment

As described in section II.C, the occupational wages of new hires are adjusted for demographic composition. This section describes how this adjustment affects the NHQI and compares the index to a version that does not use the adjustment and is instead based solely on occupational wages as found in the OES.

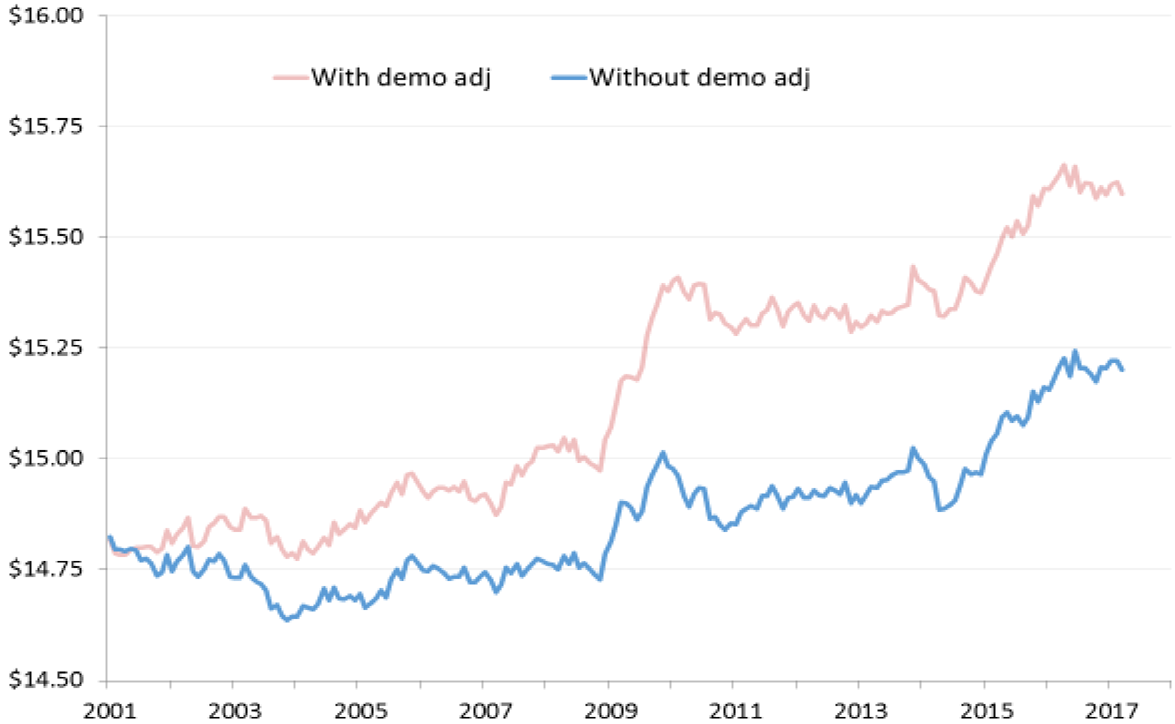
Because of the demographic adjustment, a shift in the composition of new hires toward relatively advantaged groups—the more-educated, older individuals—will tend to increase the index, while a shift toward disadvantaged groups will tend to decrease it. (The procedure allows these relationships to vary by broad occupation, but they broadly hold with varying magnitudes across occupation.)

One might be concerned that the sharp rise during the Great Recession seen in Figure 4 is driven by demographic composition of hires rather than occupational earnings power. Figure 13 investigates this possibility by comparing the NHQI wage index in levels (as it appears in Figure 4) with a version that does not apply demographic adjustment. The adjusted series is in salmon; the unadjusted series is in blue. While the two series are quite close together in 2001, at about $14.80 per hour, the adjusted series grows faster and by 2017 is about $0.40 per hour higher than the unadjusted series. The bulk of the gap between the two opened right before and during the Great Recession; since 2010 the gap has remained roughly constant.

It is thus not the case that the growth in the NHQI is solely due to changing demographic composition, as the unadjusted index is up $0.40 per hour over the entire period, and it rose by about $0.25 per hour during the Great Recession. To determine how much of the overall growth in the NHQI is due to demographic adjustment and how much is due to occupational mix, it is helpful to normalize the series to their 2005 levels. This is done in Figure 14.

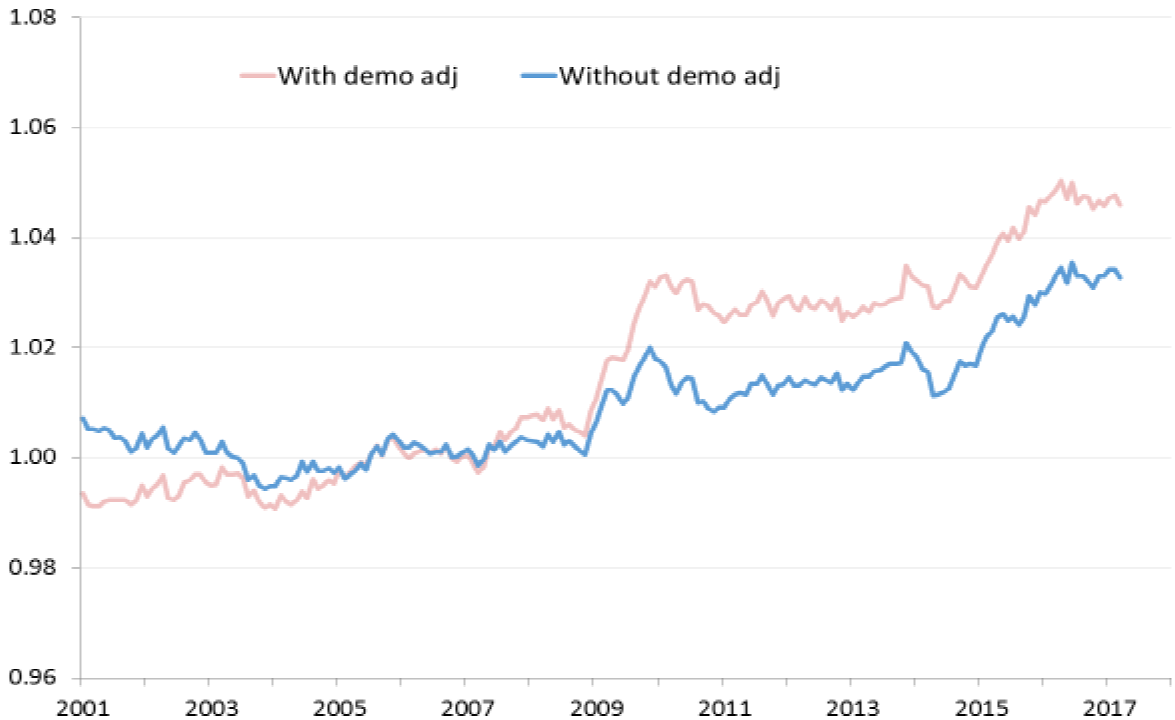
[36] Other subgroups may be added in the future, including race/ethnicity. Because of data limitations, hires per capita is available only overall, by sex and by age group.

Figure 13 NHQI: Wage Index, with and without Demographic Adjustment



SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

Figure 14 NHQI: Wage Index, by Demographic Adjustment (Normalized 2005=100)



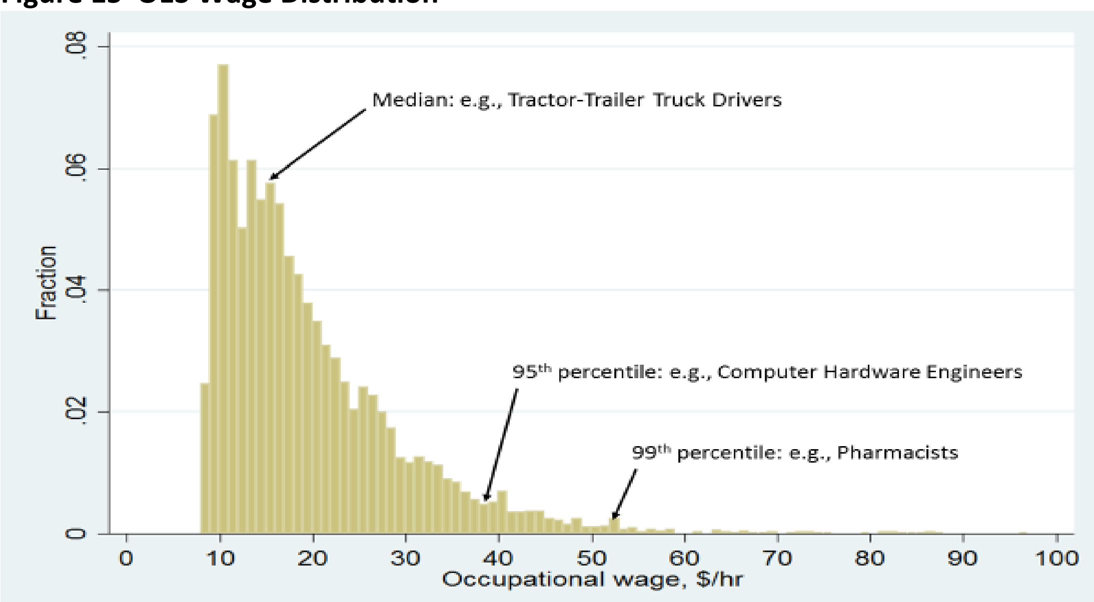
SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

Between 2005 and March of 2017, the NHQI hourly wage index grew by 4.6 percent, including demographic adjustment. Without adjustment, the index grew by 3.3 percent over the same period. Thus, about 72 percent (3.3 / 4.6) of the overall increase is due to occupation mix, and 28 percent is due to demographic adjustment. These shares vary by which starting point is used. Since the beginning of 2001, 48 percent of the increase is due to occupation mix and 52 percent is due to demographic adjustment; since 2010, the shares are instead 129 percent and −29 percent.[37] Demographics thus played a larger role in NHQI growth before the Great Recession but cannot explain the growth since.

B. Means vs. Quantiles, and Demographic Adjustment

The NHQI is calculated by taking the *mean* or average of demographically adjusted occupational wage of new hires in the CPS. (To be clear, any given individual's assigned occupational wage is taken from the 25th percentile of the OES and then demographically adjusted, and these adjusted wages are averaged together.) This sounds relatively simple, but because the occupational wage distribution is highly skewed, the process weights high-paying occupations more heavily. Figure 15 illustrates this conceptually by showing the occupational wage distribution from the OES.[38]

Figure 15 OES Wage Distribution



SOURCE: Author's tabulation from OES 2016.
NOTE: Occupational wages are 25th percentile hourly wages.

W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

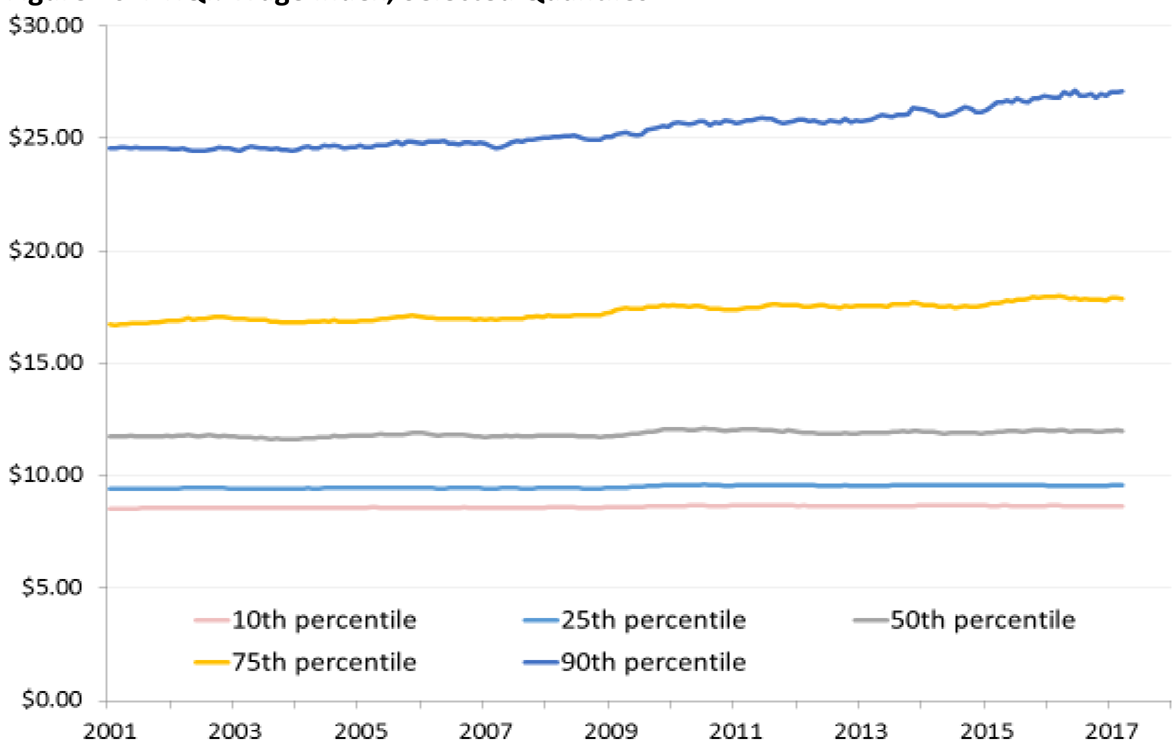
[37] This means that demographic change has worked to actually *reduce* the NHQI since 2010, and that all of the gain (and more) has come from occupational mix. Conversely, since the unadjusted index actually fell slightly between 2001 and 2009, while the adjusted index grew slightly, *all* of the NHQI growth in that period is due to demographics.

[38] The distribution represents 25th percentile wages across all six-digit SOC occupations by two-digit NAICS industries. For simplicity, the numbers have not been demographically adjusted.

The histogram shows that the majority of occupations pay less than $20 per hour (at the 25th percentile). In fact, the median or middle-paying occupation pays around $16.75 per hour, and a typical occupation at this threshold is tractor-trailer truck drivers. A few occupations, however, pay much higher amounts, as exemplified by the long right tail in the figure. Computer hardware engineers, for instance, are around the 95th percentile at $38.50, and pharmacists are roughly at the 99th percentile at $52. If new hires consist of 10 truck drivers, two computer hardware engineers, and one pharmacist (and we ignore demographic adjustment for simplicity), the average wage of the new hires would be $22.81, but the wage of the typical hire—the one earned by 10 of the 13 new hires—would still be just $16.75. Thus, a few high-paying occupations can easily shift the index.

An alternative option to taking the mean or average of adjusted wages of new hires is examining specific quantiles. This approach gives a fuller picture of where in the occupational wage distribution are changes occurring that drive the overall NHQI.

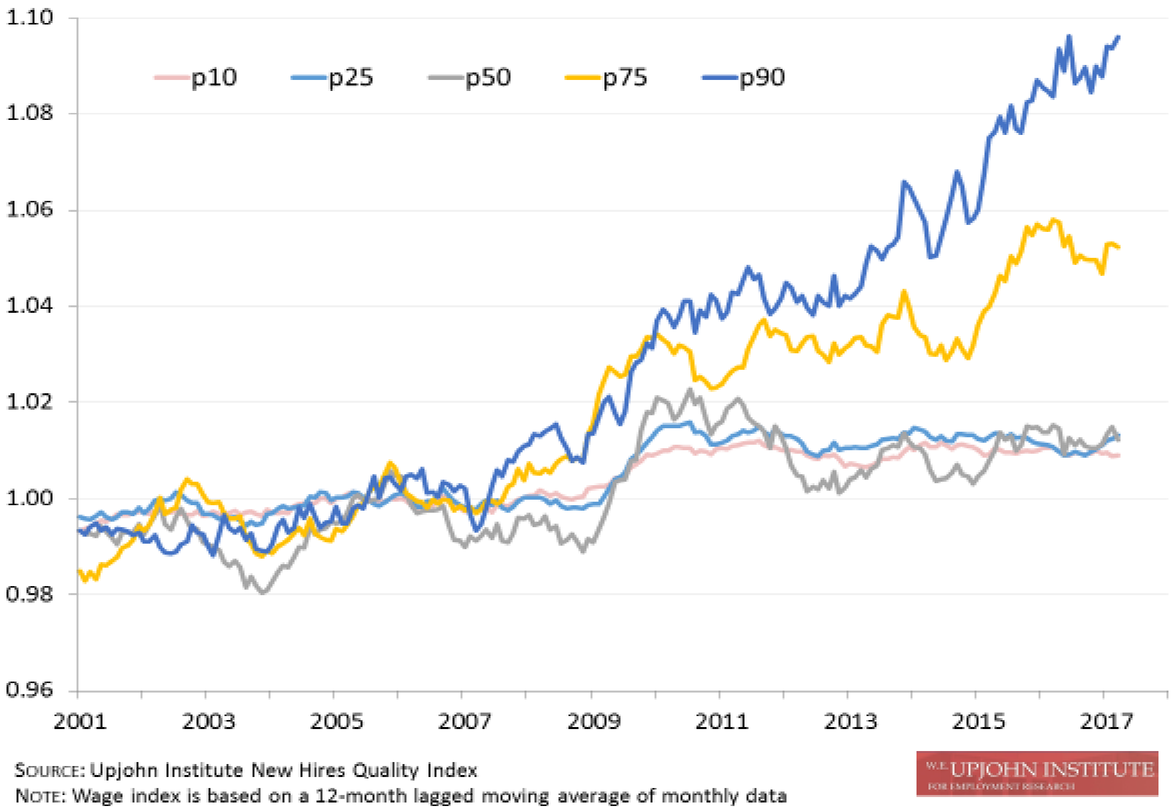
Figure 16 NHQI: Wage Index, Selected Quantiles



SOURCE: Upjohn Institute New Hires Quality Index
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

Figure 16 shows the 10th, 25th, 50th, 75th, and 90th percentiles of the NHQI in levels from 2001 through the beginning of 2017. Because the scale on the y-axis is relatively large, the lower quantiles look quite flat, although small level changes can still lead to sizable percentage changes. In line with the skewed wage distribution, the gap between the 75th and 50th percentiles is larger than the gap between the 25th and 50th percentiles, and the gap between the 90th and 75th percentiles is larger than the one between the 75th and 50th percentiles. It is also clear that—in dollar per hour terms—the largest increases in the NHQI have occurred at the higher percentiles. Of course, rather than looking at the change in levels, it can be helpful to look at proportional changes, as shown in Figure 17.

Figure 17 NHQI: Wage Index, Selected Quantiles (Normalized 2005=1.00)

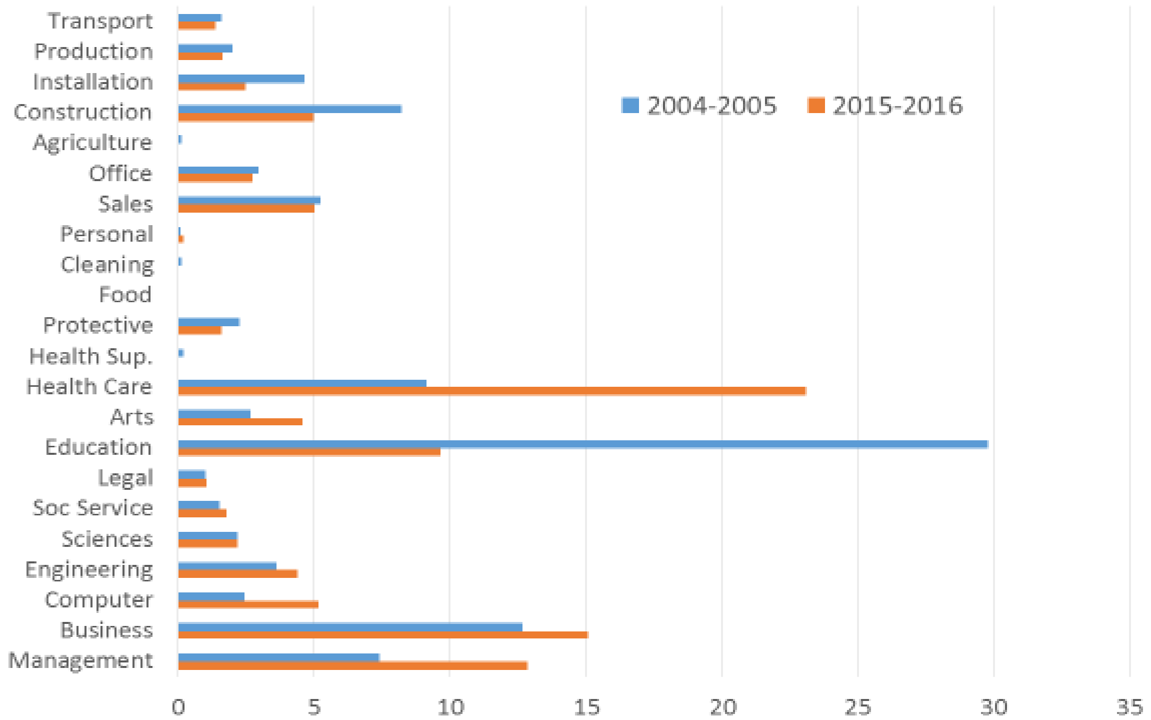


When NHQI quantiles are normalized to their 2005 levels, as shown in Figure 17, a fascinating picture emerges. First, there was little sustained trend in any of the quantiles over the 2001–2005 period, which corresponds with the relatively flat trend of the overall NHQI in this period, shown in Figure 4.

Second, the quantiles begin to diverge right around the Great Recession. At the lower end, the 10th and 25th percentiles exhibit a slight—but persistent—bump in 2009. These quantiles, which had hardly changed at all since 2001, increased by about 1 percent between 2009 and 2010 and then more or less remained constant at this higher level through the beginning of 2017. The median—the point at which half of new hires make more and half make less—is only slightly more volatile, rising between 2 and 3 percent in 2009 and 2010 before falling slightly during the recovery, and also ending at 1 percent above its 2005 level. In contrast to the pattern seen in the bottom half of the distribution, the top half, as shown by the 75th and 90th percentiles, began increasing sharply as early as 2007. For the 90th percentile, the growth seldom faltered, and by 2017 its wage index was nearly 10 percent above its 2005 level. At the 75th percentile, the initial growth between 2008 and 2010 slowed for the next five years before jumping again in 2015 and 2016, ending the period at about 5 percent above its 2005 level.

Third, these patterns imply widening inequality in who is being hired. The faster growth at the upper end is not because new hires in high-paying occupations are getting raises, but because these hires are increasingly drawn from even higher-paying (and presumably higher-skilled) occupations than in the past.[39] To make this point more concretely, Figure 18 shows the distribution of broad occupation classes that constituted 90th percentile new hires in both the 2004–2005 and 2015–2016 periods.

Figure 18 NHQI: Occupational Distribution at 90th percentile



SOURCE: Upjohn Institute New Hires Quality Index; CPS
NOTE: Data are for 89th–91st percentile of wage index for years shown.

The largest differences are in education and health care occupations. Whereas nearly 30 percent of 90th percentile jobs in 2004–2005 were in education (largely college professors and senior teachers switching jobs) and 9 percent were in health jobs, a decade later just 10 percent of such jobs were in education and 23 percent were in health. Unsurprisingly, this reflects the overall growth in health care jobs, including physicians, therapists, and advanced-practice nurses, as well as constrained budgets of many states and school districts, where much of education hiring occurs. But there is also a noticeable shift toward engineering, computer, business, and management among 90th percentile new hires, and a shift away from installation, and construction jobs. Because the former occupation groups (and health care)

[39] Commensurately, the tepid growth in the bottom half is not because these newly hired workers are not getting raises, but because the mix of occupations being hired at this part of the distribution has not shifted that much toward higher-paying (and higher-skilled) jobs.

tend to pay more than the latter groups (and education), the 90th percentile NHQI rose over time.

V. An Occupational Wage Index vs. Actual, Self-Reported Wages

The NHQI uses occupational wages provided by the OES survey in its construction, but the underlying CPS survey used to identify new hires contains self-reported wage information. Why, then, go through a complicated matching and demographic adjustment process with occupational wages rather than simply using actual, self-reported wages?

There are three problems with using the self-reported wages in the CPS. First, available sample sizes are much smaller with self-reported wages. While the CPS asks every adult individual their work status and whether they have changed employers since the previous month, it asks only one-quarter of the sample questions about wages and earnings. (Specifically, wage questions are asked only of rotation groups 4 and 8, which are known as the outgoing rotation groups, since households will not be surveyed the following month). This means that the sample size of individuals each month from which to calculate earnings power drops from 2,000–2,500 to only 500–625.

Second, many individuals refuse to answer the wage questions in the CPS, and this problem has been growing worse (Bollinger and Hirsch 2006; Bollinger et al. 2017).[40] In the late 1990s roughly one-quarter of respondents refused to answer the wage question when asked, and by 2016 this fraction had grown to two-fifths. The CPS tries to impute wages for these nonresponses (flagging observations when doing so), but the procedure is not very accurate (and even more so for new hires), and the consensus among most economists is not to use these imputations (Bollinger and Hirsch 2006). These nonresponses thus lower the effective sample size even further, to between 350 and 400 individuals with valid wage reports each month.

These reduced sample sizes significantly curtail the use of self-reported wages as a monthly index. They are too small to permit reliable calculation of a wage index for most subgroups, and even for the overall population when using a 12-month moving average, lead to much more volatile jumps over short horizons, much of which may be statistical noise. But there is also a third concern with using self-reported wages, especially when sample sizes are small: their veracity. This problem is twofold. First, among the people who report, the wage they say they earn may not capture what they truly earn, or at least what a business would report they earn on a tax form. This discrepancy arises from simple mistakes and guesswork to deliberate attempts to understate (or overstate) earnings (Bee, Mitchell, and O'Hara 2016). While new hires may be especially attuned to their wages because of recent salience, they may be more

[40] In the CPS sampling frame, one person generally reports on the characteristics of others in the household. This "proxy" response can also introduce error, especially for wages among household members who are not close relatives.

likely to deliberately misreport if the wage is especially low relative to that of a previous job. Alternatively, if the job is very high-paying and some compensation comes in the form of bonuses, the wage of a new hire may be understated. Second, people who do not report wages at all are not random but disproportionately come from the very bottom and top of the earnings distribution (Bollinger et al. 2017). Since there tends to be more churning (hiring and layoff activity) at the bottom, not counting the wages of these individuals could bias a self-reported wage index up relative to the truth.[41]

However, there are also possible advantages of actual wage data. For one, a consistent series can extend farther back in time, since there is no need to have a harmonized scheme of occupations. But another advantage stems from the conceptual difference of a wage index based on occupation, such as the NHQI, and one based on realized wages. As this report has stressed, the NHQI cannot capture wage increases within an occupation except for those due to changing demographics such as age and education. If new technology is introduced that makes software programmers more productive and these efficiency gains show up in their pay, the NHQI will not pick it up. Likewise, a minimum wage increase for retail workers also will not be included.[42] An actual wage measure, on the other hand, would capture these types of changes, and anything else that affected earnings power.

Economists often assume wages are a function of characteristics that we can measure, denoted X, how important each of these characteristics are, denoted β, and factors that remain specific to an individual, denoted ε. An actual wage measure, if there were no reporting problems, would capture all these components. The occupation-based NHQI instead captures some of the more important Xs and βs (occupation and demographics), but leaves out others potentially measurable in some data sets (e.g., location of employment, size of employer) and factors that are either vary hard to measure or that are considered specific to an individual (e.g., a person's motivation on a specific job, how good a cultural match the worker and the employer are). In this framework, an index based on actual wages plausibly better measures the economic well-being of an individual, while the NHQI is a better metric for the demand for skill among employers.

For these reasons, it can be instructive to compare the NHQI to self-reported wages in the CPS (dropping imputations but otherwise taking wages as reported). Figure 19 presents the mean and median of the NHQI (as in Figures 4 and 16, respectively), along with the means and median of self-reported hourly CPS wages.[43] Unlike previous graphs, this one extends back to 1999 to illustrate part of the longer time period possible using self-reported wages.

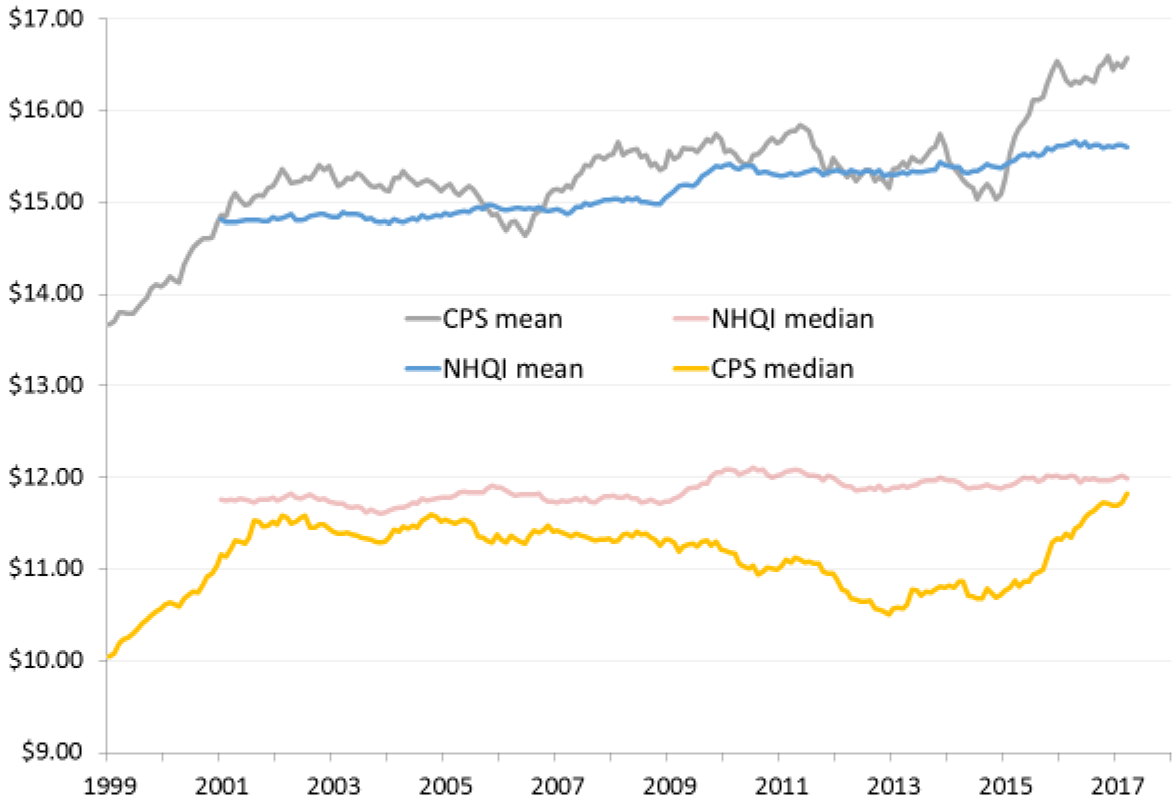
[41] Appendix Table 2 shows select characteristics of new hires compared with all incumbent workers.

[42] Similarly, declines in wages because of, for example, weakened unionization or weaker labor contracts also will not be included in the NHQI.

[43] The self-reported wages are adjusted to year 2016 dollars using the Personal Consumption Expenditures deflator from the U.S. Bureau of Economic Analysis. For individuals who report weekly wages in the CPS (salaried workers), hourly wages are weekly wages divided by usual weekly hours, which is standard practice. All series are 12-month lagged moving averages.

It turns out that the mean of self-reported hourly wages of new hires closely aligns with the mean of demographically adjusted, 25th percentile occupation wages. That is, the gray line lies near the blue line (official NHQI). This suggests that the 25th percentile is indeed a reasonable point of the occupational wage distribution to proxy for new hires. As expected given the smaller sample size and greater degree of variation among individuals than occupations, the self-reported wages series is more volatile than the official NHQI. It has also grown more, including a spurt during the well-known economic boom period in the late 1990s and early 2000s and again around 2015. In inflation-adjusted dollars, the average new hire at the beginning of 1999 earned about $13.70; 18 years later, the figure had risen to $16.50, growth of about 20 percent (or 1.04 percent per year).

Figure 19 NHQI and Self-Reported CPS Wages: Means and Medians



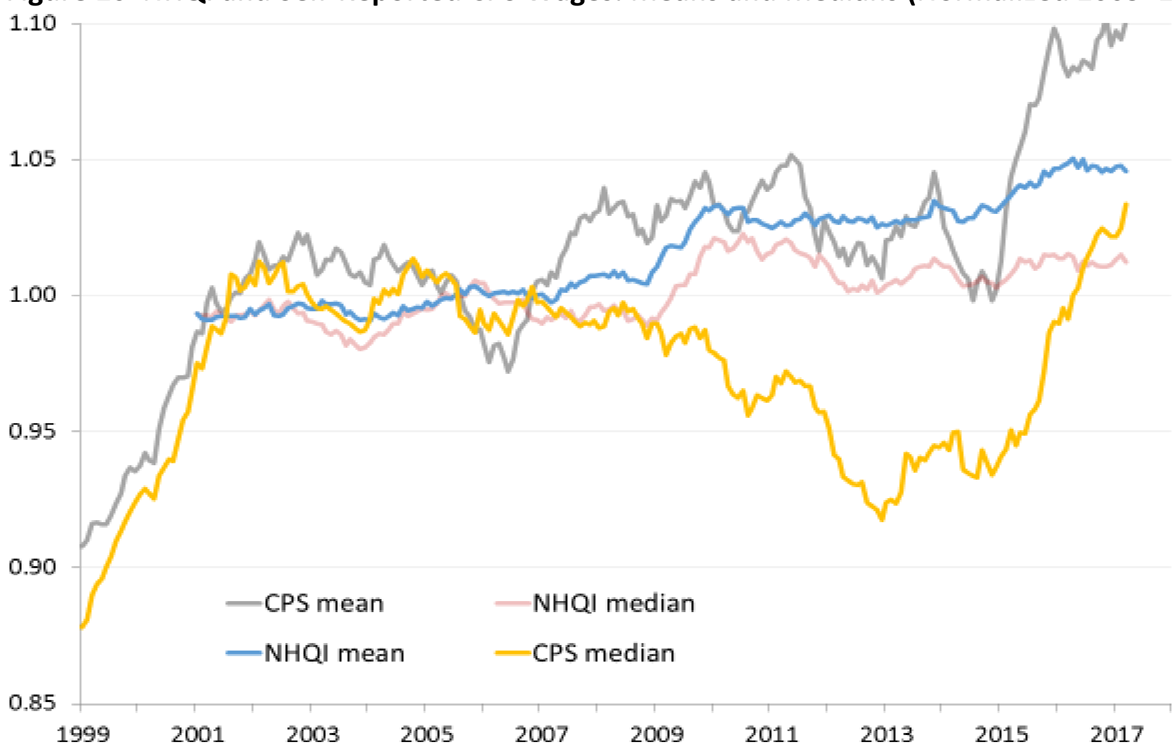
SOURCE: Upjohn Institute New Hires Quality Index; CPS
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

The two most notable discrepancies between the two series are 1) the lack of any growth in the CPS mean over the Great Recession, even as the NHQI rises, and 2) the sharp growth in 2015 in CPS wages, while the NHQI grows only modestly. For the first, the pattern suggests that actual wages were flat even as the occupations being hired became more skilled. This should not be surprising, as earnings power tends to fall during recessions. In this case, it's possible that during the Great Recession starting wages within higher-paying occupations fell, and the two effects roughly cancelled each other out, leaving a mostly flat actual wage trend. For the second, the growth of wages of new hires in 2015—which is not well known or appreciated—

occurs with minimal change in occupational skill. This suggests that the recent actual wage growth among new hires was largely within occupations rather than across them.

The bottom two lines shows the NHQI median and the median of CPS self-reported wages. Because of the skewness of the wage distribution, as explained in Figure 15, the typical new hire earns less than the "average" new hire. Remarkably, these two series are also close together. The patterns between them, however, are somewhat different than between the means. The CPS median, like the CPS mean, shows strong growth in the late 1990s and again around 2015. This implies the actual wage growth was not driven by high earners but occurred broadly throughout the distribution of new hires. Indeed, the cumulative CPS wage growth at the median is just under 19 percent, quite close to the cumulative growth at the mean. However, whereas the means of NHQI and CPS moved in tandem between 2005 and 2015 (with the slight exception around the Great Recession), the median NHQI and CPS diverge, as the latter gradually but steadily declines over this period, opening up a gap between the two series that mostly closes only over 2015 and 2016. This finding implies that among the middle of the distribution, new hire wage growth within occupations was negative for nearly a 10-year span, including before the Great Recession until several years afterward.

Figure 20 NHQI and Self-Reported CPS Wages: Means and Medians (Normalized 2005=1.00)



SOURCE: Upjohn Institute New Hires Quality Index; CPS
NOTE: Wage index is based on a 12-month lagged moving average of monthly data

Figure 20 repeats Figure 19, only normalizing each series to its level in 2005. The normalization highlights how large the relative increases in actual wages were in the late 1990s and over 2015 (and how large was the decrease in the median after the Great Recession). That recent growth in actual wages has been so robust, especially at the middle, with essentially a flat NHQI, points to strong gains in economic well-being of newly hired workers but little change in skill demand among employers.

In the future, formal decompositions between self-reported CPS wages and the NHQI may yield additional nuances into the timing of wage growth of new hires and how much of it is due to occupational and demographic factors and how much is due to other factors. This, of course, rests on the CPS self-report wages not being overly biased by the issues discussed above. For new hires in particular, it is important that the observable characteristics of those with valid self-reported wages resemble those of all new hires; otherwise, the utility of the previous comparisons will be rather limited.

A simple, if not wholly adequate, way in which to check representativeness is to compare the distribution of characteristics of all new hires to the subset that report valid wages. The approach is not wholly adequate because even if observable characteristics look similar, one cannot infer anything about other, harder to observe characteristics (such as the job motivation and cultural fit mentioned above). Nonetheless, the comparison can be illustrative in how bias may operate. Appendix Table 1 compares demographics such as age, sex, race, education, and a few others for all new hires and those with valid wages in 1999, 2007, and 2016. Several characteristics exhibit significant differences. The sample with valid self-reported wages is increasingly younger than the sample of all new hires, with this gap growing from 1.1 years in 1999 to 1.3 years in 2016. A little over one year may not sound like much, but a single year of age can affect hourly wages by 1–2 percent or more. Other major differences include race, in which the sample with valid wages consistently has more whites and fewer blacks than all new hires, and the hire type, in which the sample with valid wages increasingly has a larger share of employer changers (and thus a smaller share of newly employed) than the sample of all new hires.

To get a rough idea of how these differences can together bias a wage index based on self-reported wages, it is possible to use a regression to obtain weights for how much each of these factors matter to final wages.[44] Using just the characteristics shown in Appendix Table 1, the net effect is about a 1 percent downward bias, driven mostly by the age difference. If other characteristics are included, such as detailed industry and occupation, the bias grows somewhat, to about 2.7 percent. If this procedure is applied at different points in time, the bias appears relatively stable, as the (increasingly) negative age bias is countered by the (increasingly) positive hire type bias.

Even if this bias appears small and stable, it should provide caution in using self-reported wages to construct a wage index of new hires. The exercise cannot account for bias introduced by

[44] A regression like the one used for demographic adjustment is run, but all in one stage, using the characteristics shown in Appendix Table 1 as regressor, and possibly some others. The estimated coefficients are then multiplied by the differences in characteristics shown in Appendix Table 1, and these products are then summed to get a measure of bias.

unobservable factors, and the small sample size yields volatile oscillations that may include significant statistical noise. Still, the self-reported wages may prove complementary to the NHQI in the future.

VI. Conclusion

This report describes a new monthly index of the earnings power of newly hired workers. By identifying newly hired workers in the CPS and merging them with occupational wage data drawn from the OES and performing demographic adjustment, the Upjohn Institute's New Hires Quality Index (NHQI) provides a timely update on the quality of new job accessions. The flexibility of the NHQI leads to metrics for hourly wages, hiring volumes, new hire wage bills, and hires per capita for all newly hired workers in the country as a whole, and for more than two dozen subgroups. All these series can be updated each month with a new release of the CPS data, roughly two weeks after the Employment Situation Report comes out.

The NHQI thus complements other data releases by the federal government on new hires by industry and the wages of incumbent workers. It sheds light on whether the jobs individuals are taking are low-paying or high-paying, and how this may vary among different groups. Moreover, it provides context for how the earnings power of new hires interacts with the total quantity of new hires, adding another facet to our understanding of hiring activity. The NHQI therefore should be useful to policymakers and forecasters who desire another coincident indicator of economic activity, to journalists who seek a frequent update on the "quality" of new hires, to researchers who are interested in structural changes in skill demand, and to anyone else interested in a deeper look at job creation.

The NHQI data are available back to January of 2001 and are easily and freely accessible through an interactive web interface and for download at www.upjohn.org/nhqi.

References

Acemoglu, Daron, and David Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In *Handbook of Labor Economics: Volume 4B*, David Card and Orley Ashenfelter, eds. Amsterdam: Elsevier, pp. 1043–1171.

Bee, C. Adam, Joshua Mitchell, and Amy O’Hara. 2016. “Preliminary Research for Replacing or Supplementing the Income Question on the American Community Survey with Administrative Records.” American Community Survey Research and Evaluation Report Memorandum Series, #ACS16-RER-6. Washington, DC: U.S. Census Bureau.

Bernhardt, Annette, and Sarah Thomason. 2017. “What Do We Know About Gig Work in California? An Analysis of Independent Contracting.” Berkeley, CA: UC Berkeley Center for Labor Research and Education. http://laborcenter.berkeley.edu/what-do-we-know-about-gig-work-in-california/ (accessed August 22, 2017).

Bollinger, Christopher R., and Barry T. Hirsch. 2006. “Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching.” *Journal of Labor Economics* 24(3): 483–519.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2017. “Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch.” Mimeo.

Davis, Steven J., R. Jason Faberman, John C. Haltiwanger, and Ian Rucker. 2010. “Adjusted Estimates of Worker Flows and Job Openings in JOLTS.” In *Labor in the New Economy,* Katharine G. Abraham, James R. Spletzer, and M.J. Harper, eds. Chicago: University of Chicago Press, pp. 187–222.

Davis, Steven J., and John C. Haltiwanger. 2014. “Labor Market Fluidity and Economic Performance.” NBER Working Paper No. 20479. Cambridge, MA: National Bureau of Economic Research.

Drew, Julia A. Rivera, Sarah Flood, and John Robert Warren. 2014. “Making Full Use of the Longitudinal Design of the Current Population Survey: Methods for Linking Records Across 16 Months.” *Journal of Economic and Social Measurement* 39(3): 121–144.

Fallick, Bruce, and Charles A. Fleischman. 2004. “Employer-to-Employer Flows in the U.S. Labor Market: The Complete Picture of Gross Worker Flows.” Finance and Economics Discussion Series Working Paper No. 2004-34. Washington, DC: Federal Reserve Board.

Haltiwanger, John, Henry Hyatt, and Erika McEntarfer. 2015. “Cyclical Reallocation of Workers across Employers by Firm Size and Firm Wage.” NBER Working Paper No. 21235. Cambridge, MA: National Bureau of Economic Research.

Hershbein, Brad, and Lisa B. Kahn. 2017. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." NBER Working Paper No. 22762. Cambridge, MA: National Bureau of Economic Research.

Hyatt, Henry, and James R. Spletzer. 2013. "The Recent Decline in Employment Dynamics." *IZA Journal of Labor Economics* 2(5): 1–21.

Ihrke, David. 2014. "Reason for Moving: 2012 to 2013." P20-574. Washington, DC: U.S. Census Bureau.

Madrian, Brigitte C., and Lars J. Lefgren. 2000. "An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents." *Journal of Economic and Social Measurement* 26: 31–62.

Scopp, Thomas S. 2003. "The Relationship between the 1990 Census and Census 2000 Industry and Occupation Classification Systems." Technical Paper No. 65. Washington, DC: U.S. Census Bureau.

Appendix Table 1 Differences in Selected Characteristics between All New Hires and Those with Valid Wage Reports

	1999			*2007*			*2016*		
	All	*Valid wage*	*Difference*	*All*	*Valid wage*	*Difference*	*All*	*Valid wage*	*Difference*
Age	33.4	32.4	−1.1***	35.2	34.0	−1.2***	36.9	35.6	−1.3***
Female	0.510	0.524	0.014**	0.506	0.507	0.001	0.509	0.509	0.001
Race									
White	0.682	0.695	0.012**	0.628	0.651	0.024***	0.572	0.588	0.016**
Black	0.137	0.123	−0.014***	0.130	0.108	−0.023***	0.140	0.119	−0.021***
Asian	0.038	0.036	−0.002	0.047	0.042	−0.005*	0.058	0.056	−0.002
Hispanic	0.134	0.137	0.003	0.175	0.177	0.002	0.205	0.209	0.005
Education									
< HS	0.242	0.252	0.010*	0.212	0.219	0.006	0.165	0.163	−0.002
HS grad	0.307	0.297	−0.010*	0.300	0.296	−0.004	0.284	0.278	−0.006
Some college	0.279	0.291	0.012**	0.282	0.281	−0.002	0.304	0.318	0.014*
Bachelor’s	0.124	0.116	−0.008**	0.143	0.142	−0.001	0.165	0.162	−0.003
Grad degree	0.047	0.044	−0.003	0.062	0.062	0.000	0.082	0.079	−0.003
Region									
Northeast	0.190	0.164	−0.026***	0.195	0.178	−0.017***	0.176	0.174	−0.002
Midwest	0.234	0.239	0.006	0.219	0.233	0.014**	0.210	0.227	0.018***
South	0.341	0.349	0.008	0.362	0.356	−0.006	0.363	0.338	−0.025***
West	0.236	0.248	0.012**	0.224	0.233	0.009*	0.251	0.261	0.010
Sector									
Goods	0.214	0.209	−0.005	0.194	0.194	−0.000	0.167	0.156	−0.010*
Services	0.786	0.791	0.005	0.806	0.806	0.000	0.833	0.843	0.010
Hire type									
Newly employed	0.581	0.560	−0.021***	0.646	0.614	−0.032***	0.675	0.617	−0.058***
Changed employer	0.419	0.440	0.021***	0.354	0.386	0.032***	0.325	0.383	0.058***

SOURCE: Author’s calculations from the CPS.
NOTE: New hires are CPS longitudinal matches who transition from nonemployment to employment, or who report switching employers, between month $t - 1$ and month t. Those with valid wages are the subset of new hires who are in rotation month 4 or 8 in month t and report a (nonimputed) real hourly wage between \$2 and \$100. * p<0.10, ** p<0.05, *** p<0.01.

Appendix Table 2 Differences in Selected Characteristics between Incumbent Workers and New Hires

	1999			*2007*			*2016*		
	Incumbent	*New hires*	*Difference*	*Incumbent*	*New hires*	*Difference*	*Incumbent*	*New hires*	*Difference*
Age	39.3	33.4	−5.9***	41.0	35.2	−5.8***	42.2	36.9	−5.3***
Female	0.465	0.510	0.045***	0.464	0.506	0.042***	0.469	0.509	0.040***
Race									
White	0.744	0.682	−0.062***	0.691	0.628	−0.063***	0.640	0.572	−0.068***
Black	0.109	0.137	0.028***	0.106	0.130	0.024***	0.111	0.140	0.029***
Asian	0.038	0.038	0.000	0.049	0.047	−0.002	0.062	0.058	−0.004*
Hispanic	0.103	0.134	0.031***	0.140	0.175	0.035***	0.167	0.205	0.038***
Education									
< HS	0.126	0.242	0.116***	0.111	0.212	0.101***	0.086	0.165	0.079***
HS grad	0.316	0.307	−0.009**	0.293	0.300	0.007**	0.262	0.284	0.022***
Some college	0.284	0.279	−0.005**	0.287	0.282	−0.005**	0.289	0.304	0.015***
Bachelor's	0.184	0.124	−0.060**	0.206	0.143	−0.063***	0.230	0.165	−0.065***
Grad degree	0.090	0.047	−0.043**	0.105	0.062	−0.043***	0.133	0.082	−0.051***
Region									
Northeast	0.188	0.190	0.002	0.183	0.195	0.012	0.179	0.176	−0.003
Midwest	0.242	0.234	−0.008	0.227	0.219	−0.008	0.220	0.210	−0.010
South	0.348	0.341	−0.007	0.358	0.362	0.004	0.366	0.363	−0.003
West	0.223	0.236	0.013	0.232	0.224	−0.008	0.236	0.251	0.015
Sector									
Goods	0.247	0.214	−0.034	0.212	0.194	−0.018	0.191	0.167	−0.024
Services	0.753	0.786	0.033	0.788	0.806	0.018	0.809	0.833	0.024

SOURCE: Author's calculations from the CPS.
NOTE: New hires are CPS longitudinal matches who transition from non-employment to employment, or who report switching employers, between month *t* − *1* and month *t*. Those with valid wages are the subset of new hires who are in rotation month 4 or 8 in month *t* and report a (nonimputed) real hourly wage between $2 and $100. * p<0.10, ** p<0.05, *** p<0.01.