No Longer Qualified? Changes in the Supply and Demand for Skills within Occupations

**Mary A. Burke, Alicia Sasser Modestino,
Shahriar Sadighi, Rachel B. Sederberg, and Bledi Taska**

Abstract:

Using a novel database of 159 million online job postings, we examine changes in employer skill requirements for education and specific skillsets between 2007 and 2017. We find that upskilling—in terms of increasing demands for bachelor's degrees as well as software skills—was a persistent trend among high-skill occupations, but either a temporary or non-existent phenomenon among middle-skill and low-skill occupations. We also find evidence that persistent upskilling in the high-skill sector contributed to greater occupational mismatch that remained elevated during the recovery from the Great Recession. In contrast, labor market mismatch had largely dissipated within the low-skill and middle-skill sectors by 2017.

Mary A. Burke is a senior economist in the research department at the Federal Reserve Bank of Boston. Her email address is mary.burke@bos.frb.org. Alicia Sasser Modestino, the corresponding author, is an associate professor at Northeastern University, where she has dual appointments in the department of economics and the School of Public Policy and Urban Affairs. Her email address is a.modestino@neu.edu. Shahriar Sadighi is an economist at Amazon. His email address is shahrias@amazon.com. Rachel B. Sederberg is a visiting assistant professor of economics at Bowdoin College. Her email addresses are rachel.sederberg@me.com and r.sederbe@bowdoin.edu. Bledi Taska is the chief economist at Burning Glass Technologies. His email address is btaska@burning-glass.com.

I. Introduction

Prior research has established that US employers increased education requirements within occupations for open positions during the Great Recession, a trend that has become known as "upskilling" (Modestino, Shoag, and Ballance, forthcoming). Although roughly one-third of the upskilling that occurred during the last recession has been shown to have been cyclical or opportunistic, as much as two-thirds of the increase appears to have persisted at least through 2015 (Modestino, Shoag, and Ballance 2016; Hershbein and Kahn 2018). To the extent that upskilling has been persistent, unemployed workers in some occupations may no longer qualify for the positions they once held if they lack the necessary educational credentials to meet these new hiring requirements. To test this issue, we examine changes in employer skill and education requirements for online job openings posted in 2007 and 2010–2017—the period leading up to and following the Great Recession—and examine the implications of upskilling for labor market functioning.

Although the term "mismatch" often refers to imbalances in the supply of and the demand for labor *across* occupations, skill mismatch *within* occupations can also arise if the skill requirements for a job are changing over time. Skill mismatch within occupations could impair matching efficiency only in the affected occupations or, if the mismatch is sufficiently large or widespread, it could also reduce aggregate matching efficiency within the entire labor market. Recent evidence suggests that aggregate matching efficiency remains significantly below its pre-recession levels (Hall and Schulhofer-Wohl 2018), despite the US unemployment rate being historically low and aggregate mismatch across occupations having almost fully abated since the Great Recession (Burke 2015; Turrell et al. 2019).

Our main data source comes from Burning Glass Technologies and includes the near-universe of online job postings (roughly 159 million total) in the United States between 2007 and 2017, except for the years 2008 and 2009, when no data are available. Hence, our analysis treats 2007 as the period immediately before the onset (in December 2007) of the Great Recession—and dates the recovery from 2010 through 2017. We first use these data to determine which sectors of the labor market—classified as low-, middle-, or high-skill based on the pre-recession education levels of incumbent workers—engaged in permanent or temporary upskilling over the most recent business cycle. We find that occupations in the high-skill sector—jobs that as of 2006 were mainly held by individuals with a college degree and perhaps an advanced degree—were more likely to have experienced an increase in education requirements just after the Great Recession (2010)

1

compared with just before the recession (2007), *and* were more likely to have maintained those higher requirements between 2010 and 2017 as the labor market generally improved. In contrast, occupations in the middle- and low-skill sectors largely engaged in temporary upskilling, meaning those sectors posted higher education requirements for a given job in 2010 relative to 2007, but then reverted to less stringent requirements between 2010 and 2017 as the labor market tightened. Using the Current Population Survey, we also show that among new hires within occupations in the high-skill sector, the educational distribution shifted upwards relative to existing workers, evidence which suggests that the new education requirements imposed by employers since the Great Recession have indeed been binding.

We then use the Burning Glass data to describe changes in specific skill requirements—categorized as baseline, specialized, and software skills—within occupations in order to gain insight into what factors might have driven the changing demands for education in the three different sectors of the labor market. We find that while the demand for software skills increased rapidly within all three sectors of the US labor market between 2007 and 2010, between 2010 and 2017 demand for software skills only continued to increase in the high-skill sector. In contrast, the share of employers demanding baseline skills (e.g., project management) and other specialized skills (e.g., information security) increased in a similar manner across all sectors of the labor market—rising rapidly during the Great Recession and then continuing to rise, albeit more slowly, during the recovery. This finding is consistent with anecdotal evidence that following the Great Recession, job descriptions in some industries have been revised to include new tasks that require more advanced skills, either in response to new regulations or changes in technology. For example, new regulations under the Affordable Care Act that reduced the reimbursement rates for a host of medical procedures and services have allegedly raised the skill requirements for nurses and physician assistants, since services formerly rendered by doctors were pushed down the skill hierarchy.[1] Persistent upskilling in the manufacturing industry may be a response to new technology, such as 3D printing, which requires workers to master computer-aided design software to produce parts formerly made using analog technologies.[2] More broadly, the persistent upskilling

---

[1] See, for example, "Four Ways the ACA Affects Healthcare Staffing" at
http://www.thestaffingstream.com/2016/10/19/four-ways-the-aca-affects-healthcare-staffing/.
[2] See, for example, "2018 skills gap in manufacturing study" at
https://www2.deloitte.com/us/en/pages/manufacturing/articles/future-of-manufacturing-skills-gap-study.html.
However, these technologies may not as yet be widely used and indeed there is some evidence that employer claims of a skill gap might be overblown (Osterman and Weaver 2017).

we observe within some segments of the US labor market might represent the longer-term process known as labor market polarization, whereby trade liberalization and automation in some industries have led to a loss of or transformation in middle-skill jobs while boosting demand for both low-skill and high-skill occupations.[3]

We further identify several dimensions along which the increased demand for specialized skills and education have affected matching efficiency in the US labor market. First, since most of the persistent upskilling occurred within the high-skill sector, occupational mismatch in that sector might have moved differently between 2007 and 2017 when compared with the mismatch occurring within the middle-skill or low-skill sector. To test this hypothesis, we expand on the methodology of Şahin et al. (2014) and construct mismatch indexes by skill sector. We find that mismatch in both the low- and middle-skill sectors exhibited a sharply cyclical pattern, increasing in 2010 when compared with 2007, and then receding between 2010 and 2017. In contrast, mismatch within the high-skill sector exhibited a modest increase between 2007 and 2010 and, rather than abating, increased further between 2010 and 2017 as the labor market posted fairly steady gains.

Persistent upskilling *within* certain broad occupation groups might also have contributed to aggregate mismatch in the US labor market. To explore this possibility, we measure the extent of occupational mismatch at the level of the 3-digit Standard Occupational Classification (SOC) code, separately within each higher-level occupation (defined by 2-digit SOC codes). The idea is to test for whether the qualitative patterns in upskilling by major occupation—either on a temporary or persistent basis—predict the qualitative movements in the labor market mismatch indexes by major occupation. All of the minor occupations within several major occupations engaged in persistent upskilling, whereas other major occupations consisted mostly of minor occupations that engaged in either temporary upskilling or no upskilling between 2007 and 2017. Among the former group of major occupations, in all but one case we observe that mismatch increased between 2007 and 2010 and then continued to increase between 2010 and 2017. In contrast, for the latter group of major occupations, the mismatch indexes tended to move in concert with the aggregate unemployment rate—increasing between 2007 and 2010 and then declining (or leveling off) between 2010 and 2017. Together, these new empirical facts suggest that previous

---

[3] The literature on job market polarization is extensive: see Autor, Katz, and Kearney (2008); Acemoglu and Autor (2010); Autor and Dorn (2013), and Holzer (2015).

measures of aggregate labor market mismatch across major occupations have failed to capture labor market imbalances caused by shifts in skill demands occurring within major occupations.

Although the US unemployment rate has recently fallen to historically low levels, this drop has occurred despite the fact that aggregate matching efficiency remains below its pre-recession levels (Hall and Schulhofer-Wohl 2018). In other words, without the drag of lower matching efficiency, the US labor market might have recovered more quickly than it actually did. We identify one arguably structural trend within some occupations—employer upskilling—that may have contributed to the declines in aggregate matching efficiency observed since the Great Recession. In particular, upskilling within occupations may have contributed to skill mismatch within occupations, such that recently unemployed workers may no longer qualify for the jobs that they previously held without gaining additional training or skills. As retraining takes time, the ramping up of education and skill requirements for some jobs may have reduced matching efficiency within the affected occupations, thereby reducing aggregate matching efficiency, all else remaining equal. Furthermore, these conditions may account for the seeming contradiction, observed in the early part of the recovery period (2010–2012), that employers reported facing difficulties finding workers with the desired skills, while at the same time many economists found scant evidence of *aggregate* (rather than within-occupation) labor market mismatch.[4]

Our findings also speak to policy debates about workforce development and related educational policies, an arena which could benefit from using real-time labor market data that includes information on employer demand for specific skills. For example, the ability to distinguish between persistent versus temporary shifts in skill demands within certain occupations could help higher education institutions to identify which human capital investments are likely to have higher returns in the long run. Such information could also be used by career counselors to develop sector-based or job-driven training in key industries (as emphasized by Holzer 2015), as well as to advise job seekers about the suitability of their qualifications for various jobs and about fruitful retraining opportunities.

Our results may also help reconcile recent debates about the persistence of slow wage growth in the face of low unemployment, and employer complaints regarding the difficulty of

---

[4]Among others, Abraham (2015) and Davis, Faberman, and Haltiwanger (2012) argued against the mismatch hypothesis. Sahin et al. (2014) argued that mismatch contributed significantly to the sharp increase in unemployment during the Great Recession but that mismatch had abated significantly by early 2011.

filling vacancies. A cursory examination of wage growth by occupation indicates that wages increased rapidly between 2010 and 2017 within some of the occupations that displayed persistent upskilling, but less quickly in the occupations where upskilling was temporary—but these patterns do not hold uniformly. Yet, raising the offering wage for positions that experience a sudden upward shift in demand for skills that are in scarce supply might not be the most effective way to fill vacancies, at least in the short run when workers have not had the opportunity to seek out additional education and training. At present, there are only anecdotal stories of employers offering coding bootcamps (Lanahan 2019), but not much rigorous evidence concerning on-the-job training programs in relation to upskilling. Therefore an important question for future research concerns whether employers who engaged in permanent upskilling were also more likely to engage in training their current employees to acquire the desired skills, and whether this practice helps explains the lack of wage growth.

The paper proceeds as follows. Section II places our study within the related literature on upskilling and labor market matching efficiency. Section III describes the data and methods used to measure upskilling, matching efficiency, and labor market mismatch. Section IV presents our results. First, we examine demand-side changes in education and skill requirements within occupations over the business cycle for different skill sectors of the labor market: low, middle, and high. We then generate estimates of labor market mismatch both within these three skill sectors and within 2-digit SOC codes over the most recent business cycle. Section V concludes with a discussion of potential policy implications drawn from these findings.

II. Related Literature

Recent research suggests that changes in employer skill requirements reflect a combination of both cyclical and structural forces. On the cyclical side, Modestino, Shoag, and Ballance (forthcoming) show that the share of job postings requiring a four-year college degree increased by 10 percentage points during the Great Recession. They estimate that about one-third of the upskilling they observed was an opportunistic response to the greater availability of workers during that period. In a separate paper, these same authors show that employer demand for college degrees, as well as other types of skills, fell as the labor market tightened between 2010 and 2014 (Modestino, Shoag, and Balance 2016).

On the structural side, a complementary set of papers shows that up to two-thirds of the

5

upskilling that occurred during the Great Recession was in fact persistent, meaning that it was apparently structural in nature and reflected a permanent change in job requirements (Hershbein and Kahn 2018, Zago 2018). Structural upskilling may also be related to the longer-term trends of skill-biased technological change (Katz and Murphy 1992; Autor, Katz, and Krueger 1998; Autor, Levy, and Murnane 2003) and labor market polarization (Autor, Katz, and Kearney 2008; Autor and Dorn 2008; Acemoglu and Autor 2010). In addition, both cyclical forces and the structural forces stemming from the Great Recession may be reinforcing these trends, as recessions have been shown to induce long-term changes in the labor market (Hershbein and Kahn 2018; Charles, Hurst, and Notowidigdo 2012; Jaimovich and Siu 2012; Tuzeman and Willis 2013; Beaudry, Green, and Sand 2013).

Perhaps not coincidentally, the Beveridge curve exhibited a significant outward shift during the Great Recession, such that the unemployment rate appeared elevated relative to what was expected at the given vacancy rate, a condition that persisted through at least 2012. Since then, a large literature has developed that tries to account for this shift in the Beveridge curve. These explanations have included the role of reduced employer recruiting efforts (Davis, Faberman, and Haltiwanger 2012), increased uncertainty (Barnichon et al. 2012; Daly et al. 2012), pre-recession trends in matching efficiency (Hall and Schulhofer-Wohl 2018), extended unemployment benefits (Daly et al. 2012; Veracierto 2011; Barnichon and Figura 2010; Hagedorn et al. 2014), and cyclical fluctuations in job search effort (Mukoyama, Patterson, and Şahin 2018), among other factors.

One remaining explanation that has been repeatedly put forth by employers, but which has received very little support from the empirical literature, is labor market mismatch. The one notable exception is Şahin et al. (2014), which showed that the mismatch between job vacancies and unemployed workers across industries, occupations, and geographies increased significantly during the Great Recession, contributing to lower matching efficiency in the aggregate, and therefore helped account for the shift in the Beveridge curve. However, the sluggish wage growth observed during most of the recovery period seems inconsistent with the mismatch hypothesis (Rothstein 2012; Abraham 2015). Several papers have argued that weak aggregate demand offered a more convincing explanation for the outward shift of the Beveridge curve than either skills mismatch or other structural factors (Barlevy 2011; Lazear and Spletzer 2012; Rothwell 2012; Carnevale, Javasundera, and Cheah 2012; Diamond 2013; Diamond and Şahin 2015; Osterman and Weaver 2017). Furthermore, by 2014 measures of labor market mismatch across occupations

had also recovered to pre-recession levels (Burke 2015). Nonetheless, recent evidence suggests that aggregate matching efficiency remained below its pre-recession level even after 2015 (Hobijn and Perkowski 2016; Hall and Schulhofer-Wohl 2018).

For the recent US labor market experience, starting from the Great Recession and going through the ensuing recovery, relatively little work has been done to study the relevance of skills mismatch *within* occupations, and there are very few formal studies of changes in skill requirements by occupation. Using the US Department of Labor's O*NET database, Vaisey (2006) compares the current education requirements for various jobs (as reported in surveys rather than in job ads) to the educational attainment of incumbent workers employed in these same jobs, and finds that the average worker was overqualified for his/her job as of 2002. Liu and Grusky (2013) also find evidence that certain skill requirements measured using O*NET—including computer skills, analytical skills, and quantitative skills—increased within job categories since 1979, but that the increases have ranged from small to modest. Although skill-biased technological change has been cited as a factor leading to increased demand for highly-educated workers relative to less-educated workers (Katz and Murphy 1992; Autor, Katz, and Krueger 1998; Autor, Levy, and Murnane 2003), the evidence is mixed as to whether the adoption of new technologies raises skill requirements *within* jobs (for example, see Acemoglu 2002; Zicklin 1987; and Keefe 1990). More recently, Hershbein and Kahn (2018) have found that the persistent upskilling in job requirements that was observed during the Great Recession emerged as the result of routine-biased technological change that was occasioned by the recession itself.

Some researchers have questioned whether the requirements listed on job openings are truly binding, or rather are indicative of problems related to overqualification, rather than underqualification, within occupations (Cappelli 2014 and Cappelli 2015). However, jobs that require relatively high computer usage, including occupations in the fields of science and healthcare, experienced large increases in wages at the upper end of the wage distribution relative to occupations involving less computer usage (Bessen 2014). Similarly, states that experienced greater job polarization during the recession—defined as a loss of routine middle-skill jobs compared to low-skill (manual) and high-skill jobs—also experienced a greater mismatch in educational qualifications during the recession and through the end of 2013, as evidenced by workers moving down the occupational skill ladder relative to what would be expected given their higher educational attainment (Zago 2018). Over the longer run, structural labor market

polarization may manifest as upskilling within some occupations, as employers raise skill requirements, and as educational overqualification in other job categories, as former middle-skill workers are forced to take jobs in the low-skill sector.

We build on these different strands in the labor market literature by describing patterns in upskilling that have taken place over a longer time period, and decomposing these trends by skill sector, according to whether the job vacancies initially tended to require either low-, middle-, or high-education levels. We further identify which types of occupations engaged in structural versus temporary upskilling in their education requirements. By using the richness of the data from Burning Glass Technologies (BGT), we then examine which types of underlying skills (baseline, specialized, or software) might be driving the large and persistent increases in the demand for more education observed in some job listings. While exploring these trends is of inherent interest, we make an important contribution to the literature by examining the implications of these changes in skill demands. We find that over the business cycle, mismatch evolved very differently in the high-skill sector compared with the middle-skill and low-skill sectors, a result that aligns with the observed differences in upskilling patterns. We also find evidence that mismatch exhibited secular (rather than cyclical) increases within certain major occupations that exhibited persistent upskilling. Taken together, these two new empirical facts suggest that persistent upskilling in the high-skill sector of the US labor market may have exerted a drag on aggregate matching efficiency during the recovery from the Great Recession. Furthermore, this source of drag has not been detected in most previous approaches to the measuring labor market mismatch.

III. Data and Methods

Our primary objective is to explore the degree to which the upskilling that occurred within occupations during the current business cycle contributed to reduced matching efficiency in specific portions of the US labor market and possibly in the aggregate labor market. We seek to answer these specific research questions:

- How have employers' minimum education requirements increased within occupations and in which skill segments (low, middle, and high) of the US labor market? Are these higher skill requirements binding?
- Which skillsets are now in greater demand? Are these increases persistent or temporary?
- Does upskilling generate mismatch—within and/or across occupations? If so, does this help explain the reduced matching efficiency in certain segments of the labor market?

To test our hypotheses, we use a variety of methods drawing on multiple data sources to provide a variety of evidence from which we draw our conclusions. Our analysis primarily consists of two parts. First, we use the near-universe of online job vacancy data provided by BGT to examine the demand-side changes in education and skill requirements within occupations over the business cycle, describing such changes separately for the three major skill sectors (low, middle, and high). Second, we generate separate estimates of the occupational mismatch between 2007 and 2017 for each of these skill sectors and for several major occupations in order to identify linkages between upskilling and labor market mismatch. Along the way, we also conduct several robustness checks to confirm whether these higher skill requirements appear to be binding and explore the pattern in wage increases across occupations engaged in persistent versus temporary upskilling.

A. Data Sources

1. Online Job Posting Data from Burning Glass Technologies

Over the past two decades, online vacancy data have been used by a number of researchers to study labor market dynamics (Kuhn and Skuterud 2004; Bagues and Labini 2009; Şahin et al. 2014; Marinescu 2017). The vacancy data used in this paper is collected by BGT, one of the leading vendors of online job posting data. BGT collects detailed information on the more than 7 million current online job openings in the United States posted daily from over 40,000 sources that include job boards, newspapers, government agencies, and employer websites.[5] The data are collected via a web crawling technique used to browse online job boards and other websites, then systematically text parse the wording in each job listing into usable data elements. From free-text job postings, BGT mines over 70 job characteristics, including employer name, location, job title, occupation, years of experience, and level of education required.[6]

Unlike other vendors of online job posting data, BGT also parses out other skill dimensions from the advertisement's text, information that allows us to categorize the sought-for job skills precisely and directly, rather than simply relying on educational requirements as a proxy for the required work-related skills. BGT aggregates these data by parsing each skill listed in a job posting

---

[5] See http://www.burning-glass.com/realtime/ for more details.
[6] Note that the BGT data do not contain any information on the duration of the vacancy, how many applications a job listing received, or whether a vacancy was filled.

and categorizing the skill into a canonical version of similarly named skills. For example, Python 3.3 and Python 2.7 are both standardized to Python. Because this process results in over 16,000 canonical skills, BGT then places each canonical skill into a broader skill cluster. For example, algebra and calculus would both be placed into the math skill cluster. These skill clusters are then aggregated upwards into skill families (e.g. math and science are both STEM skills). Finally, skill families are classified as either baseline skills (e.g. leadership), specialized skills (e.g., accounting), or software skills (e.g., Oracle).

BGT's data-collection process is designed to capture the most current and complete set of online postings at a given time and includes algorithms that eliminate duplicate ads for the same job vacancy. To avoid large fluctuations over time, BGT places more weight on large job boards that aggregate job listings, rather than on individual employer sites, which are updated less frequently.[7] BGT uses the same filtering and de-duplication algorithm across years and applies any improvements made to the algorithm retroactively, so the methodology used is consistent throughout our dataset. Nonetheless, the number of sources scraped by BGT may have evolved over time.

The advantage of using online job vacancy data is that it allows analysis at a greater frequency and at more refined geographic locations than what is possible to achieve when using traditional employer surveys, such as the Job Openings and Labor Turnover Survey (JOLTS). The BGT data are available at a monthly frequency by zip code and at the level of the six-digit Standard Occupation Classification (SOC) code for 2007 and 2010–2017.[8] One potential drawback is that the online vacancy data only capture job listings posted on the internet, so therefore may not be representative of the universe of available job openings if vacancies from certain industries and occupations are less likely to be posted electronically. However, estimates show that as of 2012, between 60 and 70 percent of all US job vacancies were posted online (Carnevale, Jayasundera, and Repnikov 2014).[9] Previous research shows that the number of online job advertisements represents a reliable predictor of actual hiring activity one quarter later (Templin and Hirsch 2013),

---

[7] BGT has also provided access to their Labor/Insight analytical tool that enables us to access the underlying job postings to validate many of the important components of this data source including timeframes, de-duplication, and aggregation.

[8] No data are available from BGT for 2008 and 2009.

[9] Carnevale, Jayasundera, and Repnikov (2014) audited a sample of job postings in the BGT database and found that the BGT coding for occupation, education, and experience was accurate at least 80 percent of the time. Given the algorithm improvements that have been retroactively and consistently applied since then, this 80 percent accuracy rate is likely an underestimate.

and other authors have also tested the robustness of these data (Rothwell 2014; Hershbein and Kahn 2018). We explore this issue in the appendix. Despite differences in the sampling frame of the BGT data when compared to state and national employer job vacancy surveys, we find that the industry and occupation distributions of the different vacancy series are quite similar and that the relationships between the BGT data and the survey-based series are consistent over time.[10]

We make use of two distinct versions of the BGT data. The main dataset is provided at the job-posting level and contains detailed data on the job title, occupation, industry, and location, as well as information on the requirements for education, experience, and skills. The data are pooled over the given year without duplication, and we observe the month in which the posting first appeared. We use this version of the data for the first part of our analysis to explore the changes over time in employer demands by education and detailed types of skills at the 3-digit and 6-digit SOC code level.

Although Figure 1 shows that the BGT data exhibit a high degree of correlation over time with the JOLTS vacancy series (0.82), the level of the vacancies in the BGT data is consistently lower. In part, this is because the JOLTS data include both online and offline job listings, such as those job openings that are circulated in print media, posted in shop windows, or posted internally within firms. In addition, the JOLTS specifically asks about the number of vacancies that are open, whereas a single online job posting may represent multiple vacancies. To obtain an accurate number of vacancies over time, BGT creates a normalized (reweighted) dataset that exactly matches the monthly JOLTS vacancies by industry. We use this second version of the data when calculating the number of vacancies by occupation on the demand side to construct our mismatch indices by skill sector.

2. Current Population Survey Microdata

On the labor supply side, we make extensive use of the Bureau of Labor Statistics's Current Population Survey (CPS) microdata (Flood et al. 2018) from 2007 through 2017. To construct our mismatch indexes, we use the CPS data to measure the number of unemployed workers by occupation at various levels of occupational detail (such as the 3-digit SOC code and the 6-digit

---

[10] These comparisons show that the BGT data tend to slightly over-represent industries such as finance, and Slightly under-represent some other industries, such as food services. Similarly, occupations such as management are slightly over-represented, while occupations such as food preparation are slightly under-represented.

SOC code). When estimating the amount of unemployment related to labor market mismatch, we exploit the panel dimension of the CPS data to estimate job finding rates and job destruction rates by occupation, utilizing the methodology developed by Shimer (2012). Finally, we make further use of the longitudinal data in the CPS to identify new hires by the 6-digit SOC code and continuing employees by the 6-digit SOC code. We also use the CPS data to describe the education level of new hires relative to continuing employees by occupational skill sector, and to make similar educational comparisons between other pairs of types, such as continuing employees versus unemployed workers. The details of how we construct these various measures from the CPS data are provided below and in the appendix.

B. Methods

1. Classifying Occupations by Skill Sector

Our data analysis partitions the universe of occupations into three skill sectors—low, middle, and high—and examines the labor market patterns (such as upskilling or mismatch) for each skill sector. An occupation is classified into one of these three sectors according to the pre-recession distribution of educational attainment among incumbent workers, as measured by the 2005–2007 sample of the American Community Survey (ACS). An occupation is classified as "high-skill" if at least a plurality (40 percent) of individuals employed in that line of work had completed at least a four-year college degree, while an occupation is considered "low-skill" if at least 40 percent of its workers had no more than a high school education. Occupations that satisfied neither of these criteria—that is, occupations in which fewer than 40 percent of employees had four-year degrees *and* fewer than 40 percent had only high school education or less—are classified as "middle-skill."[11] These three skill definitions remain fixed throughout the period being studied, 2007 through 2017.

2. Measuring Demand-Side Upskilling within Occupations by Skill Sector

We use the BGT data to examine the demand-side changes in education and skill requirements within occupations that occurred between 2007 and 2017 for each of the three major

---

[11] Note that we observe no cases in which an occupation could be placed into more than one skill category   based on these criteria.

occupational skill sectors: high, low, and middle. We are interested in whether the increases in average job requirements by occupation (or by skill sector) were either temporary—defined as being restricted to changes on either side of the Great Recession (between 2007 and 2010)—or instead persisted into the recovery period, defined as running from 2010 through 2017. We define an occupation as having upskilled across the Great Recession if between 2007 and 2010 the share of postings in the occupation requiring a four-year college degree or higher increased by at least 5 percentage points—this threshold represents the average increase in the share of job postings requiring a four-year college degree during this period, as shown in Modestino, Shoag, and Ballance (forthcoming). Occupations for which the share of job openings that required a four-year college degree increased by less than 5 percentage points between 2007 and 2010 are classified as having "no upskilling." An occupation is defined as having experienced "persistent upskilling" if, having first upskilled between 2007 and 2010, the share of job postings requiring a four-year college degree declined by less than 2.5 percentage points between 2010 and 2017. In contrast, an occupation that upskilled between 2007 and 2010 is defined as having experienced "temporary upskilling" if the share of job postings requiring a four-year college degree reverted by more than 2.5 percentage-points during the recovery period. Using these definitions we classify each 3-digit SOC code as having experienced either no upskilling, temporary upskilling, or persistent upskilling. We then characterize upskilling behavior at the level of the skill sector (high, middle, or low), based on the respective shares of 3-digit occupations in the sector that experienced each type of upskilling.

Besides examining general changes in education requirements, we also examine changes in requirements for specific skills, as these may serve as indicators of structural changes in the nature of the underlying job. The BGT data contain detailed descriptions of job skills appearing in the online vacancy ads, and these are grouped into 16,000 canonicalized skill clusters, which in turn are aggregated into three broad categories of skills—"baseline skills," "specialized skills," and "software skills." Baseline skills refer to general skills that cut across many different occupations, such as communication and familiarity with Microsoft Office. Specialized skills refer to narrower skills that are required only for some types of jobs, including teaching, litigation, and accounting. Software skills include specialized programming skills, such as knowledge of Java Script, SQL, and Oracle. Given these categories, we use a difference-in-difference analysis to test for whether occupations that engaged in persistent upskilling also tended to exhibit larger and/or

more persistent increases between 2007 and 2017 in the share of postings requiring skills of a given type (e.g., baseline, specialized, or software skills). We then delve further into the skills data, describing changes in the most frequently requested skill clusters within each of the baseline, specialized, and software skill categories. This analysis seeks to reveal whether employers were simply seeking more of the same skillsets for a given occupation or requiring an entirely new set of skills.

3. Measuring Labor Market Mismatch by Skill Sector

Heterogeneity in upskilling behavior across the three occupational skill sectors may have differential implications for matching efficiency within each skill sector. For example, if persistent upskilling is more likely to occur within the high-skill sector compared to the low- or middle-skill sectors, then we would expect the high-skill sector to exhibit different movements in mismatch over the business cycle—across the recession and/or during the recovery period—compared with sectors that experienced temporary upskilling or no upskilling. One way to detect such differences is to measure labor market mismatch separately by skill sector over the period of interest, as labor market mismatch offers one indication of inefficiencies in the hiring process. To do so, we extend the methodology of Şahin et al. (2014) to construct separate occupational mismatch indexes for each of our three skill sectors (low, middle, and high).

In this framework, the labor market is frictional in the sense that a given worker is assumed to search for jobs only within a circumscribed set of ocupations (in this case, within a 2-digit SOC code restricted by skill sector). Likewise, this model assumes that firms only hire workers who are searching within the occupational classification that corresponds to the job opening. These assumptions seem more plausible for major occupational categories (such as the level of the 2- or 3-digit SOC code) than for more narrowly defined and detailed occupational categories (e.g., the level of a 6-digit SOC code). Although some studies show that mobility across occupations is quite low, these estimates have been known to vary widely.[12]

Following Şahin et al. (2014), we adopt a labor market mismatch index that quantifies the fraction of potential hiring that does not take place because of an occupational misallocation of

---

[12] For example, Molloy, Smith, and Wozniak. (2017) observed that during the 2000s, the CPS showed that there was a 4 percent transition rate across three-digit SOC codes, while Kambourov and Manovskii (2009) reported finding a 21 percent transition rate across 3-digit SOC codes during the 1990s using the Panel Study on Income Dynamics.

unemployed workers relative to the distribution of vacancies by occupation.[13] Here we provide a brief derivation of the mismatch index, with further details presented in the appendix. The number of hires in occupation i at time t, denoted as h_{it}, is assumed to be governed by a matching process that can be represented as:

$$h_{it} = \phi_t \varphi_{it} m(u_{it}, v_{it}) = \phi_t \varphi_{it} u_{it}^{1-\delta} v_{it}^{\delta} \ . \tag{1}$$

In the above equation, $\phi_t \varphi_{it}$ represents matching efficiency for occupation i at time t, which includes both an aggregate component, ϕ_t, and an occupation-specific component, φ_{it}. Throughout our mismatch calculations, these parameters are held constant. The expresssion $m(u_{it}, v_{it})$ represents an underlying matching function that follows a Cobb-Douglas form, where hires are increasing both in the number of unemployed workers searching for a job in occupation i at time t, denoted as u_{it}, and in the number of vacancies for occupation i at time t, denoted by v_{it}. Summing across all occupations (and assuming a common matching function across them), the aggregate number of hires in the economy, h_t, can be written as:

$$h_t = \phi_t u_t^{1-\delta} v_t^{\delta} \left[\sum_{i-1}^{I} \varphi_{it} \left(\frac{u_{it}}{u_t} \right)^{1-\delta} \left(\frac{v_{it}}{v_t} \right)^{\delta} \right]. \tag{2}$$

In the above expression, u_t and v_t refer to the total number of unemployed workers and the total number of vacancies, respectively. Assuming unemployed workers could be assigned by a hypothetical social planner to search for jobs in any occupational labor market, and treating the vectors of occupation-specific vacancies and matching efficiencies as given, the planner would assign more unemployed workers to search for jobs in occupations with more vacancies and/or higher matching efficiencies, up to the point at which the total number of hires is maximized.[14] Using the expression above, together with the expression for the maximal number of hires (derived in the appendix), the basic mismatch index can be written as follows:

$$M_t = 1 - \frac{h_t}{h_t^*} = 1 - \sum_{i=1}^{I} \left(\frac{\varphi_{it}}{\bar{\varphi}_t} \right) \left(\frac{v_{it}}{v_t} \right)^{\delta} \left(\frac{u_{it}}{u_t} \right)^{1-\delta} \ , \tag{3}$$

---

[13] Labor market mismatch can also be measured across industries. Throughout our discussion, "industries" may be substituted for "occupations" with no loss of generality.

[14] If there were no differences in matching efficiency across occupational labor markets, then in order to maximize hires, the share of unemployed workers assigned to search for jobs in a given occupation should exactly match the share of total job openings accounted for by openings in that occupation.

where $\bar{\varphi}_t$ refers to a CES aggregator of the occupation-specific matching efficiencies weighted by their respective vacancy shares. The mismatch index captures the total number of hires that fail to occur under the given (inefficient) allocation of unemployed workers to occupational labor markets, as a fraction of the optimal number of hires.

We expand on the basic (i.e., aggregate) mismatch index by measuring occupational mismatch separately by skill sector. To construct these indexes, we partition each major 2-digit SOC code into three bins—a low-skill bin, a middle-skill bin, and a high-skill bin—based on the educational categories of each of the six-digit SOC codes within the 2-digit code.[15] We make similar partitions for the demand side (the number of vacancies) and the supply side (the number of unemployed). Then the mismatch index for each sector is constructed based on equation (3), which requires inputting the respective vectors of the vacancy shares and unemployment shares by 2-digit SOC code by skill sector. For example, within the high-skill sector, the vacancy share for a given (skill-partitioned) 2-digit SOC code represents the total number of vacancies *in the high-skill bin* within that 2-digit code as a share of all high-skill vacancies, and similarly for the unemployment share of a given (skill-partitioned) 2-digit SOC code. If we observe elevated mismatch within the high-skill sector, that would mean that unemployed individuals in the high-skill sector are searching in the wrong high-skilled occupations (at the 2-digit SOC level) relative to the distribution of vacancies for high-skilled occupations (at the 2-digit SOC level).[16]

By making some further assumptions, we can use the mismatch index to develop an estimate of "mismatch unemployment," defined by Şahin et al. (2014) as the difference between actual unemployment and the level of unemployment that would arise if there were no mismatch. To calculate the counterfactual unemployment rate in the absence of the occupational mismatch, we follow Şahin et al. (2014) and Shimer (2005). See the online appendix for further details regarding the estimation of mismatch unemployment by sector.

---

[15] For example, we partition 29, the 2-digit SOC code for healthcare practitioners and technical occupations, into each 6-digit category within the 2-digit designation. We then categorize each 6-digit occupation as low skill, middle skill, or high skill using the prior definitions based on the 2005–2007 ACS. For example, ob-gyn (29–1064) is categorized as a high-skill occupation, pharmacist (29–1051) is categorized as a middle-skill occupation, and OSHA technician (29–9012) is categorized as a low-skill occupation. We do this for all the 6-digit occupations within healthcare practitioners and technical occupations (29). We then create separate 2-digit healthcare practitioners and technical occupations for each skill sector by aggregating the appropriate 6-digit occupations within given high-, middle-, or low-skill sector. See Figure A5 in the appendix.

[16] Note that unemployed individuals "belong" to the high-skill sector if their most recent employment spell was in a high-skill 6-digit occupation, and they are assumed to search only within the high-skill bin of the 2-digit occupation containing their most recent 6-digit occupation. See Figure A.6 in the appendix.

4. Estimating Mismatch within the 2-Digit SOC Codes

Upskilling could also occur through a shifting of vacancies *within* a major 2-digit SOC code toward jobs requiring more education and more specialized skills. Under the assumptions of the mismatch model, if workers cannot easily move across the 3-digit occupations within the 2-digit SOC category, then persistent upskilling would lead to high and persistent measures of mismatch. To test for this possibility, we construct mismatch indexes across the minor 3-digit occupations within each of the major 2-digit occupation groups, subject to data limitations.[17] If structural upskilling leads to an increase in mismatch within an occupation, then the mismatch index should exhibit less cyclicality, and be steady or increasing during the recovery period for those 2-digit occupations that exhibited a high prevalence of persistent upskilling. We determine the prevalence of persistent upskilling for a 2-digit SOC code according to the share of 3-digit SOC codes under its umbrella that engaged in persistent upskilling.

IV. RESULTS

A. Demand-Side Upskilling within Occupational Skill Sectors

We know from previous research that employers increased skill requirements within occupations during the Great Recession—most notably, by requiring a four-year college degree or higher. About one-third of this upskilling was a cyclical response to the increased supply of workers, a trend that reversed as the US labor market recovered (Modestino, Shoag, and Ballance 2016). However, recent research finds that as much as two-thirds of the upskilling that occurred during the recession was structural (Hershbein and Kahn 2018).

1. Educational Upskilling within Occupations by Skill Sector

In this subsection, we exploit the detailed information in the BGT data to gain a better understanding of the heterogeneity in upskilling by skill sector that occurred between 2007 and 2017. Across all US occupations, just under 40 percent exhibit persistent upskilling, close to 19 percent show evidence of temporary upskilling, and almost 42 percent engaged in no upskilling (see Panel A of Table 1). However, persistent upskilling was far more prevalent within the high-skill occupation group (82.1 percent), compared to 40.0 percent in the middle-skill sector and 11.6

---

[17] For some 3-digit occupations with small numbers of employees and low rates of unemployment, the CPS contained insufficient data to generate the supply side of the mismatch index.

percent in the low-skill group. In contrast, temporary upskilling was most common among middle-skill occupations (25.0 percent), while 69.8 percent of low-skill occupations experienced no upskilling). Splitting the data differently, Panel B of Table 1 shows that high-skill occupations were sharply overrepresented among the set of all occupations that experienced persistent upskilling—comprising close to 64 percent of that group, whereas low-skill occupations made up 79 percent of the set of non-upskilling occupations.

How did these upskilling patterns evolve within each skill sector over the business cycle? Panel A of Figure 2 shows the aggregate trend in education requirements for the low-, middle-, and high-skill sectors. Again, we find that persistent upskilling was more prevalent in the high-skill sector of US the labor market. Between 2007 and 2010, the share of job postings requiring a four-year college degree increased across all three skill groups, although much more so for middle- and high-skill occupations. Between 2010 and 2017, the share of postings requiring a four-year degree declined for low- and middle-skill occupations, but remained elevated and or even increased among high-skill occupations.

Although the skill and/or education requirements listed in the BGT job postings data are an indicator of employer demand, the advertisements may represent aspirational qualities for job candidates, not strictly binding qualifications. If employers were even partly successful in upskilling their jobs—in the sense of hiring a higher share of college graduates into a given occupation—we would expect the average education level of new hires to increase relative to that of continuing employees. Since persistent upskilling was more prevalent within the high-skill sector, we would expect any increase in the relative education level of new hires to be most pronounced within that sector.

As a robustness check to test whether increasing education requirements were in fact binding, we use the longitudinal component of the CPS to plot the ratio of the average education level of new hires to that of continuing employees by skill sector for the period from 2005 through 2017. Among high-skill occupations, where much of the persistent upskilling occurred, we find that the relative average education level among new hires displayed a modestly increasing trend over the time period, notwithstanding some fluctuations (see Figure A.7 in the appendix). In contrast, among middle-skill occupations, where much of the temporary upskilling occurred, this same ratio increased sharply after 2008 and remained elevated through 2014, but then reverted back to its pre-recession levels (or lower) by 2015. These facts suggest that employers in the

middle-skill sector hired more-educated workers only as a temporary phenomenon. In the low-skill sector, where most occupations experienced no significant upskilling, the relative education level of new workers over continuing employees did not change significantly between 2005 and 2017. These differential patterns suggest that the increased education requirements reflected more than mere aspirations, and that employers were at least partly successful in meeting their demand for more highly educated employees.

2. Changes in Required Skill Clusters within Occupations by Skill Sector

While these trends are suggestive, increases in education requirements among those occupations that experienced persistent upskilling may not reflect structural changes in the actual skills required in order to do the job, but instead may have occurred in response to a permanent increase in the supply of workers with four-year college degrees over the time period (as opposed to a temporary decline in the reservation wage of such workers during the recession). To test for actual changes in skill (rather than just education) requirements, Table 2 performs a difference-in-difference analysis of changes in the education and skill requirements for persistent versus temporary upskilling occupations over time—first considering changes on either side of the Great Recession and then looking at changes during the recovery period. The top row shows that the share of postings requiring a four-year college degree increased significantly between 2007 and 2010 within both the persistent- and temporary-upskilling groups (although by a larger margin in the former group). However, between 2010 and 2017, the persistent upskilling group exhibited a further increase in the share requiring a completed college degree—by nearly 4 percentage points—compared to a drop of 12.6 percentage points in the temporary-upskilling group. This exercise confirms that the differential changes in education requirements by skill sector since 2010 were in fact statistically significant, as found in Modestino et al. (forthcoming).

Even more striking is that a similar pattern holds when performing a difference-in-difference analysis of measurable skill requirements taken from the individual job postings. For this analysis we partition the set of specific skill requirements into three categories—baseline skills (e.g., communication), specialized skills (e.g., accounting), and software skills (e.g., Oracle)—and then calculate, for each occupational skill sector (low, middle, or high) and each skill category (baseline, specialized, or software), the change over a specific time period (2007–2010 or 2010–2017) in the share of job openings in the given skill sector that required skills in the given category.

19

The remaining rows in Table 2 show that both for persistent- and temporary-upskilling occupations, the share of job postings specifying each of these three types of skills increased between 2007 and 2010. However, for any given skill category (e.g. specialized skills), the increase between 2007 and 2010 in employer demand for the skills (as derived from the vacancies) does not differ significantly between occupations that exhibited temporary upskilling and those that exhibited persistent upskilling. During the entire recovery period (2010–2017), both persistent and temporary upskilling occupations showed further increases in the share of postings requiring baseline and specialized skills, although the increase is statistically significant only among occupations that engaged in persistent upskilling. In contrast, also during the recovery, the share of postings requiring software skills was flat for jobs that displayed temporary upskilling but continued to increase among persistent upskilling occupations. Summing up the results, we observe a statistically unambiguous difference between temporary-upskilling occupations and persistent-upskilling occupations in terms of the changes over time that each displayed in the demand for software skills, whereas the time patterns in skill demands for the other skill categories are harder to distinguish based on upskilling behavior. These results suggest that the persistent increases in education requirements may have been complementary to increased demand for software skills.

It is also useful to describe the changes between 2007 and 2017 in the demand for different skillsets when examined separately by occupational skill sector, rather than by comparing skill demand trends based on upskilling behavior. Panel B of Figure 2 shows the time trends in skill requirements within each of the low-, middle-, and high-skill sectors. Each of the three occupational skill sectors experienced a sharp increase in the share of postings requiring baseline and specialized skills between 2007 and 2010, followed by some amount of reversion in demand for these same skillsets as the labor market started to recover (between 2010 and 2014), and yet each sector also saw a second uptick in skill demands as the labor market tightened further from 2014 through 2017. In contrast, the demand for software skills evolved somewhat differently depending on the skill sector. Between 2007 and 2010 there was an increase in the demand for software skills in all three skill sectors, but the increase was smaller for low-skilled occupations compared with the increases seen in the middle- and high-skill sectors. More importantly, between 2010 and 2017, the share of postings requiring software skills was relatively flat for low-skill occupations but increased significantly among both the middle-skill and high-skill occupations.

These patterns reinforce the previous finding that time trends in the demand for software skills differed between groups of occupations based on their upskilling behavior, because persistent upskilling was more prevalent among middle-skill and (especially) high-skill occupations than among low-skill occupations.

To obtain more precise information about the nature of the skills listed in job postings for the different skill sectors in the US labor market and on how these skill demands changed between 2010 and 2017, we divide each of the three broad skill categories (baseline, specialized, and software) into a set of more narrowly defined skill clusters. Figure 3, which pertains only to skill clusters within the baseline skills group, plots the share of job ads (for a given skill sector) that as of 2010 required a particular skill cluster against the change between 2010 and 2017 (for the same skill sector) in the share of postings that required that same skill cluster. Figures 4 and 5, respectively, plot the analogous information for skill clusters within the specialized skills group and the software skills group. These figures reveal differences in the demand for detailed skillsets across the three skill sectors, and provide insights into which skillsets saw the largest increases in demand between 2010 and 2017.

Looking across the three different panels of Figure 3, some baseline skill clusters appear in job ads within all three of the occupational skill sectors, such as communication and computer literacy. Nonetheless, we observe clear differences across the skill sectors regarding the extent of the initial demand in 2010 for the various baseline skills, and these differences align with our prior beliefs, indicating that the data are meaningful. For example, high-skill occupations are more likely to require research, planning, writing, and problem-solving abilities when compared with middle-skill occupations, while middle-skill occupations are more likely to require organizational skills when compared with low-skill occupations, which in turn are more likely to require physical abilities (such as strength) when compared to middle-skill and high-skill occupations. Within each skill sector (corresponding to a single panel in Figure 3), a simple regression of the change between 2010 and 2017 in the share of job ads requiring a given skill against the share of jobs requiring that skill in 2010 yields a positive slope, suggesting that the skills (within a given set of occupations) that were more commonly listed in job ads as of 2010 also experienced larger increases in prevalence between 2010 and 2017.

Figure 4 shows that patterns in the demand for "specialized" skill clusters also match our intuition about the types of specialized skills associated with different types of occupations. For

example, high-skill occupations are more likely to require teaching, budget management, and business strategy than do other occupations. In contrast, middle-skill occupations are more likely to require scheduling and retail industry knowledge, while low-skill occupations are more likely to require experience in food and beverage service, equipment repair and maintenance, and material handling. The respective regression lines fitted to the data in each panel display positive slopes, indicating that within each class of occupations the specialized skill clusters that were initially more common also tended to experience larger increases in prevalence between 2010 and 2017. However, on average that relationship appears weaker than it did in Figure 3, based on the fact that the actual data points in Figure 4 exhibit larger deviations from the fitted regression lines, especially for the plots pertaining to the middle-skill and low-skill occupations.

As shown in Figure 5, and consistent with our general findings concerning software skills, patterns in the demand for software skill clusters differ sharply across the low-, middle,-and high-skill sectors.[18] Middle-skill jobs and low-skill jobs are much less likely to require software skills than are high-skill jobs. Among middle-skill postings from 2010, the most common software skill cluster required was Microsoft Office, and even that fairly basic skill was observed in only 0.7 percent of postings. Among low-skill jobs, less than 0.3 percent of postings listed any individual software skillset. In contrast, high-skill job openings, not surprisingly, are much more likely to require advanced software skills, such as SQL programming, statistical software, and Oracle, and increases in software skill demands between 2010 and 2017 were more pronounced among the high-skill jobs compared with jobs in the other two skill sectors. Also, unlike the patterns we observed for baseline skills and specialized skills, between 2010 and 2017 employers in high-skill occupations increasingly sought applicants with specific software skills that were less commonly listed in job ads as of 2010, such as skills in integrated development environments, rather than simply seeking more of the same skills over time. These findings are consistent with our earlier evidence that advances in technology may be driving the persistent upskilling in high-skill occupations, particularly those that use specialized software packages (e.g., architectural design) or for which new software rapidly diffuses, thus changing the nature of the worker's core tasks. For example, beginning in 2014 the Affordable Care Act mandated the transition to electronic health records, and as a result health care providers were required to learn specialized programs

---

[18] Note that due to the greater heterogeneity in the software skills that are requested, in our data the prevalence of any particular software skill is much lower than that of a particular baseline or specialized skill.

such as Epic recordkeeping software (Johnson 2016).

B. Trends in Labor Market Mismatch

1. Mismatch across Occupations by Skill Sector

Based on the evidence presented in the previous sections, we hypothesize that the persistent and possibly structural increases in skill requirements that were concentrated within high-skill occupations between 2007 and 2017 may have led this sector to experience a greater persistence in labor market mismatch during the recovery when compared to middle- and low-skill occupations. Specifically, the upskilling trend suggests that changes in the distribution of US job vacancies within the high-skill sector favored occupations requiring more education and advanced skills. For a given distribution of unemployed workers by occupational experience, the shifting vacancy composition could have resulted in increases in occupational mismatch. The temporary upskilling that occurred in the low-skill and middle-skill sectors might have produced different mismatch trends across the business cycle (2007–2017) when compared to the high-skill sector. Of course, we acknowledge that upskilling was not the only factor that might have influenced occupational mismatch during the period we observe, but the thesis that persistent upskilling took place in some sectors of the labor market does have the potential to explain the persistent weakness in the US labor market that helped prolong the recovery.

To test these hypotheses, we extend the methodology of Şahin et al. (2014) to construct separate mismatch indexes (across major occupations) within each of the low-, middle-, and high-skill sectors. Panel A of Figure 6 plots the mismatch index for each sector at a monthly frequency for 2007 and 2010 through 2017 (the years for which the normalized BGT data are available).[19] The figure reveals several interesting patterns. First, the *level* of mismatch is increasing in the education level of the sector: the mismatch index is highest for the high-skill occupations (0.2 to 0.25) and lowest for low-skill occupations (0.04 to 0.075). This finding suggests that while having more education makes workers more adaptive, it also makes them more specialized and hence less substitutable across occupational categories— this second effect seems to dominate, even at the 2-

---

[19] As a robustness check, we also replicate and extend the aggregate mismatch index from Şahin et al. (2014) across all occupations and industries using the BGT data and get very similar results to their study, which used the Conference Board's Help Wanted OnLine vacancy data. Results are available from the authors upon request.

digit SOC level.[20] For example, an individual with a four-year college degree or higher who works in the Architectural and Engineering occupational group likely will be unable to switch costlessly to a job in another two-digit occupational group—even one that is somewhat related, such as the Computer and Mathematical group.

Second, *changes over time* in our measure of mismatch also vary by skill sector. In the high-skill sector, mismatch fluctuated moderately over the time period under consideration: the index increased slightly between January 2007 and December 2007, then receded between 2010 and 2012, only to increase again through mid-2014, and then stayed roughly constant or softened slightly through 2017. On balance, the mismatch index changed very little between 2007 and 2016–2017, but the increase during part of the recovery, from 2012 through mid-2014, indicates that cyclical demand fluctuations were not driving these changes. In contrast, among middle-skill occupations the mismatch index increased relatively sharply within 2007 and then fully reversed these gains on net between January 2010 and December 2017, notwithstanding the fact that the index increased moderately between January 2013 and mid-2016. These movements are consistent with a temporary run-up in skill requirements for middle-skill jobs between 2007 and 2010 that subsequently receded between 2010 and 2017. Labor market mismatch in the low-skill sector exhibited an even more strongly countercyclical pattern, increasing sharply within 2007, then falling steadily during the recovery and reverting to roughly its initial pre-recession value by the end of 2017. The temporarily elevated mismatch in the middle-skill sector may also be reflective of job market polarization, a phenomenon that has included the disappearance of routine and manual job opportunities in the middle-skill occupations (Acemoglu and Autor 2011), particularly during recessions (Jaimovich and Siu 2014).

Panel B of Figure 6 shows estimates by skill sector of the extent to which mismatch contributed to the US unemployment rate, at a monthly frequency, between January 2007 and December 2017 (not including any months in 2008 or 2009). These estimates refer to the difference between the actual US unemployment rate (by sector and month) and the counterfactual unemployment rate that would have occurred in that sector had there been no mismatch. Among high-skill occupations, the contribution of mismatch to the unemployment rate peaked at 1.6

---

[20] Şahin et al. (2014) performed a similar, but not identical, exercise that appears in their online appendix. They also found that occupational mismatch was increasing given the average education level of the incumbent workers in the set of occupations under consideration.

percentage points in January 2012. Despite subsequently receding, as of December 2017 this contribution was about 0.9 percentage point, which remained elevated relative to its pre-recession range (between 0.45 and 0.8). In contrast, labor market mismatch in the middle-skill sector contributed a maximum of just 1.0 percentage point to unemployment in that sector (in January 2010), and as of December 2017 its contribution to unemployment was, at just under 0.4 percentage point, squarely within the range of its pre-recession (January 2007 through December 2007) values. Similarly, labor market mismatch in the low-skill sector contributed at most roughly 1.0 percentage point to the unemployment rate (in mid-2011), and as of December 2017 contributed only about 0.45 percentage points to unemployment, at the low end of the range of its contributions observed between January and December of 2007.

How much of the persistent weakness in the US labor market during the recovery from the Great Recession can be explained by mismatch, and how does this contribution differ by skill sector? Using the absolute contributions of mismatch to the unemployment rate by sector shown in Figure 6 (Panel B), we calculate the *share* of the actual unemployment rate (by sector and year) that can be attributed to mismatch. (We average this share over all the months within each year to get the annual values.) Table 3 shows the resulting estimates from 2010 and 2015. For reference, the table also shows the absolute contribution of mismatch to unemployment by sector and year. Among high-skill occupations the share of unemployment due to mismatch *increased* significantly, from 31 percent in 2010 to 47 percent in 2015. In contrast, among low-skill and middle-skill occupations the share of unemployment due to mismatch declined slightly between 2010 and 2015.

2. Mismatch Within Occupations by the Prevalence of Upskilling

Based on our findings, reported above, that a significant share of high-skill occupations engaged in persistent upskilling between 2007 and 2017, in this subsection we consider whether such upskilling led to occupational mismatch ***within*** major occupational groupings. The idea is that within a major occupational classification, upskilling could have led vacancies to shift to more specialized minor occupations that had relatively few experienced workers to draw on for filling these open positions. To explore this concept, we calculate separate mismatch indexes for major 2-digit SOC codes based on the 3-digit SOC minor occupations that fall under its umbrella. Because we are slicing the BGT data more finely, we limit our analysis to 2-digit SOC codes that

had at least 1 million open positions posted online in 2010, and focus on the recovery period running from 2010 through 2017.

We focus on comparing qualitative patterns of within-occupation mismatch across different major occupations, according to their upskilling behavior. First, we examine mismatch patterns within each of three major (2-digit SOC) high-skill occupations characterized by "pure upskilling"—defined as a major occupation in which all of its minor (3-digit SOC) occupations engaged in persistent upskilling between 2007 and 2017. Figure 7 shows that for two out of three of these cases (e.g., Management; Computer and Mathematical)—we observe that mismatch over the recovery period either remained roughly flat or experienced a net increase since 2010—as we would expect if persistent upskilling affected matching efficiency. The one exception was Business and Financial occupations, which displayed a countercyclical mismatch pattern. In contrast to these three high-skill occupations, Figure 8 shows that mismatch within each of the three major middle-skill occupations—those in which upskilling was either temporary or absent—was flat or declining, as would be expected when the labor market is recovering. Finally, Figure 9 shows that mismatch within each of the nine low-skill occupations that experienced little or no upskilling during the Great Recession also exhibited patterns that were flat or declining during the recovery period. Overall, 14 out of the 15 cases studied are consistent with the theory that that persistent upskilling contributed to occupational mismatch that was either stable or rising during the recovery period, findings that contrast with the more common pattern of a countercyclical increase in upskilling followed by an abatement during the recovery. Clearly, however, there are other factors at play that can affect matching efficiency, as demonstrated by our one exception, occupations in business and finance.

The evidence that we have presented thus far suggests that there is some connection between the persistent rise in employer skill requirements and stable or rising occupational mismatch, a potential contributor to reduced matching efficiency in the aggregate labor market and therefore a potential explanation of the outward shift in the Beveridge curve.[21] However, the

---

[21] We estimate the Beveridge curve by skill sector and find that the slope of the Beveridge curve is correlated with education, meaning that the high-skill occupations exhibit the steepest relationship and low-skill occupations have the flattest relationship. In addition, most of the improvement in the Beveridge curve has come from movements in the curve for low- and middle-skill occupations, which both show large reductions in unemployment as the number of vacancies increased. In contrast, the reduction in unemployment among high-skill occupations has been much smaller relative to the number of vacancies created. While this is suggestive evidence of the persistent "wedge" that economists have observed in the aggregate Beveridge curve during this period, other factors affecting vacancy yields

sluggish wage growth observed during most of the recovery period seems inconsistent with the mismatch hypothesis (Rothstein 2012; Abraham 2015). To explore this further, we take a cursory look at wage growth by 2-digit SOC code across occupations based on their upskilling behavior. We find that six out of the 10 persistent upskilling occupations—and three out of five of the "pure upskilling" occupations—experienced wage growth between 2010 and 2015 that was faster than in the aggregate (see Table A.3 in the appendix). Within the non-upskilling group, only two out of nine occupations had above-average wage growth. Moreover, most of the five occupations characterized as having experienced "pure upskilling" also experienced employment growth during this period that exceeded the aggregate US employment growth, but this was not the case for the non-upskilling occupations. Prior research suggests that if occupational mismatch is present, then employment growth will be positively correlated with wage growth, a finding that indicates that the employment and wage growth trends reported here are consistent with the presence of mismatch in occupations that exhibited persistent upskilling in the high-skill sector (Abraham 2015).[22]

V. Conclusions and Policy Implications

Using a novel database of roughly 159 million online US job postings aggregated by Burning Glass Technologies, we find that movements in the demand for and supply of skills between 2007 and 2017—a period that included the Great Recession and the ensuing recovery period—vary across low-, middle-, and high-skill sectors. On the demand side of the labor market, education requirements increased across all occupations between 2007 and 2010 but stayed elevated between 2010 and 2017 only among high-skill occupations. In contrast, middle-skill occupations exhibited more temporary upskilling—with increases in education requirements between 2007 and 2010 that were in most cases reversed between 2010 and 2017—and the low-skill sector experienced little to no upskilling between 2007 and 2017. Comparing trends in requirements for actual skill categories reveals that the demand for software skills increased in a persistent manner (through

---

and matching efficiency may also be shifting the Beveridge curve during this period. See Figures A.12 though A.14 in the appendix.

[22] One reason for this result could be that mismatch arises when some occupations have very high vacancy to unemployment (V/U) ratios while other occupations have very low ones. The occupations with high V/U ratios will need to raise wages significantly to attract more employees, since their markets are very tight, whereas the occupations with low V/U ratios will have low employment growth (low demand) since they can fill their vacancies by offering smaller wage increases.

2017 rather than merely through 2010) only among occupations that exhibited persistent upskilling, a phenomenon which was largely confined to the high-skill sector.

On the supply side of the labor market, the education levels of new hires in the high-skill sector increased relative to the education levels of continuing employees between 2005 and 2017, suggesting that employers succeeded in hiring increasingly qualified workers. In contrast, the education level of new hires relative to continuing employees increased only temporarily within the middle-skill sector, consistent with the view that the temporary upskilling in that sector was related to the temporary increase in the supply of highly-educated workers during the Great Recession that eventually subsided as the recovery progressed. In the low-skill sector, which experienced no significant upskilling between 2007 and 2017, we correspondingly observe no significant changes in the education level of new hires relative to continuing employees over the same time period.

We also find evidence from a variety of sources that upskilling contributed to reduced matching efficiency in portions of the US labor market, either in a temporary or a persistent fashion, during the recovery. For example, patterns in occupational mismatch since 2010 differed across the three skill sectors in a way that aligns with the patterns in upskilling observed by sector: high-skill occupations consistently displayed higher mismatch rates than did middle-skill or low-skill occupations, and mismatch in the high-skill sector increased mid-recovery without fully reverting. In contrast, mismatch in the middle-skill and low-skill sectors moved in a strongly countercyclical fashion, recovering to very low levels by 2017. Considering mismatch-induced unemployment by skill sector, the share of the unemployment rate in the high-skill sector induced by mismatch actually increased between 2010 and 2015, whereas the corresponding share in either the low-skill or the middle-skill sector decreased over the same time period. In addition, mismatch trends measured within major occupational groupings differ according to the extent of upskilling among minor occupations in the umbrella group. Taken together, these exercises suggest that lower matching efficiency in the high-skill sector of the US labor market may reflect a shift in demand towards more specialized jobs, thus leading to imbalances between the demand for and the supply of certain skills.

We offer some caveats to the conclusions that can be drawn from our findings. Our mismatch index and our mismatch unemployment estimates are based on the numbers of *unemployed* job-seekers in various occupations. These measures ignore how other types of job-

seekers—such as employed people conducting a search and individuals counted as non-participants in the labor force—might influence the potential hiring rate and therefore the assessment of matching efficiency at any point in time. Şahin et al. (2014) showed that their estimates of occupational mismatch were robust to including on-the-job searchers. On the other hand, Hall and Schulhofer-Wohl (2015) find that estimates of matching efficiency depend significantly on whether job-to-job seekers and non-labor-force job-seekers are included.

Labor supply-side factors might explain why we do not always observe a persistent elevation in mismatch within occupations that experienced persistent upskilling, nor do we always observe dramatic increases in wages where we would expect such rises to occur. Just as there is evidence that temporary upskilling was an opportunistic phenomenon used by employers to take advantage of slack labor demand during the Great Recession, long-run increases in the supply of highly educated workers may also have enabled persistent upskilling to occur, apart from the temporary supply increase that resulted from the downturn.

Our findings have important implications for the economics literature as well as for labor market policy. In terms of the empirical literature, our results identify an additional factor, and plausibly a structural one, that contributed to reduced aggregate matching efficiency but which has not been investigated in detail before now. Furthermore, our results suggest that search-and-matching models of the labor market, in which unemployed workers accumulate specific human capital and make explicit mobility decisions across distinct labor markets, may need to account for the fact that workers could be chasing a moving target—at least when examining the labor market for high-skilled occupations (Kambourov and Manovskii 2009; Alvarez and Shimer 2011; and Carrillo-Tudela and Visscher 2013). Going forward, perhaps these frameworks could be modified to investigate the dynamics that lead job-seekers to search for work in the wrong sectors.

In terms of the implications for policymaking, the characteristics of those occupations experiencing more persistent shifts in skill and education requirements may point to the potential structural forces underlying these observed trends. To this end, our findings can be used to inform workforce development strategies and related educational policies, where decisionmakers could benefit from the use of real-time information on the demand side of the labor market to provide guidance for both job placement and program development geared to meeting employer demands for skills.

References

Abraham, Katharine G. 2015. "Is Skill Mismatch Impeding U.S. Economic Recovery?" *ILR Review* 68(2): 291–313.

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature* 40(1): 7–72.

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks, and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, David Card and Orley Ashenfelter, eds. Volume 4a, 1043–1071. Amsterdam: North Holland Publishing.

Alvarez, Fernando and Robert Shimer. 2011. "Search and Rest Unemployment." *Econometrica* 79(1): 75–122.

Autor, David H. 2014. "Skills, Education and the Rise of Earnings Inequality among the 'Other 99 Percent.'" *Science* 344(6186): 843–851.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103(5): 1553-1597.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90(2): 300–323.

Autor, David H., Lawrence F. Katz, and Alan B. Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113(4): 1169–1213.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118(4): 1279–1334.

Bagues, Manuel F., and Mauro S. Labini. 2009. "Do On-line Labor Market Intermediaries Matter? The Impact of Almalaurea on the University-to-Work Transition." In *Studies of Labor Market Intermediation*, David H. Autor, ed. 127–154. Chicago: University of Chicago Press.

Barlevy, Gadi. 2011. "Evaluating the Role of Labor Market Mismatch in Rising Unemployment." *Economic* 35(Q3): 82–96.

Barnichon, Regis, and Andrew Figura. 2010. "What Drives Movements in the Unemployment Rate? A Decomposition of the Beveridge Curve." Finance and Economics Discussion Series, Number 2010-48, Divisions of Research & Statistics and Monetary Affairs.Washington, DC: Board of Governors of the Federal Reserve System.

Barnichon, Regis, Michael Elsby, Bart Hobijn, and Ayşegül Şahin. 2012. "Which Industries are Shifting the Beveridge Curve?" *Monthly Labor Review*, June.

Beaudry, Paul, David A. Green, and Benjamin M. Sand. 2013. "The Great Reversal in the Demand for Skill and Cognitive Tasks." Working Paper 18901. Cambridge, MA: National Bureau of Economic Research.

Bessen, James. 2014. "Employers Aren't Just Whining—the Skills Gap is Real." *Harvard Business Review*, August 25.

Burke, Mary A. 2015. "The Rhode Island Labor Market in Recovery: Where is the Skills Gap?" Federal Reserve Bank of Boston: Current Policy Perspectives No. 15-7.

Cappelli, Peter. 2014. "Skill Gaps, Skill Shortages, and Skill Mismatches: Evidence for the US." Working Paper 20382. Cambridge, MA: National Bureau of Economic Research.

Cappelli, Peter. 2015. "Skill Gaps, Skill Shortages, and Skill Mismatches: Evidence and Arguments for the United States." *ILR Review* 68(2): 251–290.

Carnevale, Anthony P., Tamara Jayasundera, and Ban Cheah. 2012. "The College Advantage: Weathering the Economic Storm." Washington, DC: Georgetown University Center on Education and the Workforce.

Carnevale, Anthony P., Tamara Jayasundera, and Dimitri Repnikov. 2014. "Understanding Online Job Ads Data: A Technical Report." Washington, DC: Georgetown University Center on Education and the Workforce.

Carrillo-Tudela, Carlos, and Ludo Visscher. 2013. "Unemployment and Endogenous Reallocation Over the Business Cycle." CESifo Working Paper Series No. 4079.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. 2012. "The Masking of the Decline in Manufacturing Employment by the Housing Bubble." *Journal of Economic Perspectives* 30(2): 179-200.

Daly, Mary, Bart Hobijn, Ayşegül Şahin, and Robert Valletta. 2012. "Search and Matching Approach to Labor Markets: Did the Natural Rate of Unemployment Rise?" *Journal of Economic Perspectives* 26(3):3–26.

Davis, Steven J., R. Jason Faberman, and John C. Haltiwanger. 2012. "Recruiting Intensity during and after the Great Recession: National and Industry Evidence." *American Economic Review* 102(3): 584–588.

Diamond, Peter A. 2013. "Cyclical Unemployment, Structural Unemployment." *IMF Economic Review* 61(3): 410–455.

Diamond, Peter A., and Ayşegül Şahin. 2015. "Shifts in the Beveridge Curve." *Research in Economics,* 69(1): 18–25.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles and J. Robert Warren. 2018. Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D030.V6.0

Hagedorn, Marcus, Fatih Karahan, Iourii Manovskii, and Kurt Mitman. 2014. "Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects." Staff Reports No. 646. New York: Federal Reserve Bank of New York.

Hall, Robert E., and Sam Schulhofer-Wohl, 2018. "Measuring Job-Finding Rates and Matching Efficiency with Heterogeneous Jobseekers." *American Economic Journal* 10(1): 1-32.

Hershbein, Brad. 2017. "The New Hires Quality Index: A Wage Metric for Newly Hired Workers." Staff Paper, October 11. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. Available at: https://upjohn.org/nhqi/reports/NHQI_report.pdf.

Hershbein, Brad, and Lisa Kahn. 2018. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." *American Economic Review* 108(7): 1737–1772.

Hobijn, Bart, and Patryk Perkowski. 2016. "The Industry-Occupation Mix of U.S. Job Openings and Hires." Working Paper. Available at http://dx.doi.org/10.2139/ssrn.2858603.

Holzer, Harry J. 2015. "Job Market Polarization and U.S. Worker Skills: A Tale of Two Middles." Working Paper. Washington, DC: The Brookings Institution. Available at https://www.brookings.edu/wp-content/uploads/2016/06/polarization_jobs_policy_holzer.pdf.

Jaimovich, Nir, and Henry E. Siu. 2012. "Job Polarization and Jobless Recoveries." Working Paper No. 18334. Cambridge, MA: National Bureau of Economic Research.

Johnson III, Ralph J. 2016. "A Comprehensive Review of an Electronic Health Record System Soon to Assume Market Ascendancy: EPIC®." *Journal of Healthcare Communications* 2016, 1:4. DOI: 10.4172/2472-1654.100036

Kambourov, Gueorgui, and Iourii Manovskii. "Occupational Mobility and Wage Inequality." *Review of Economic Studies* 76(2): 731–759.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1): 35–78.

Keefe, Jeffrey H. 1990. "Numerically Controlled Machine Tools and Worker Skills. *Industrial and Labor Relations Review* 44(3): 503.

Kuhn, Peter J., and Mikal Skuterud. 2004. "Internet Job Search and Unemployment Durations." *American Economic Review* 94(1): 218–232.

Lanahan, Lawrence. 2019. "What If We Hired for Skills, Not Degrees?" The Hechinger Report. https://hechingerreport.org/what-if-we-hired-for-skills-not-degrees/ May 4, 2019)

Lazear, Edward P., and James R. Spletzer. 2012. "The United States Labor Market: Status Quo or New Normal?" Economic Policy Symposium – Jackson Hole, Federal Reserve Bank of Kansas City, pp. 405-451.

Liu, Yujia, and David B. Grusky. 2013. "The Payoff to Skill in the Third Industrial Revolution." *American Journal of Sociology* 118 (5): 1330–1374.

Madrian, Brigitte C., and Lars John Lefgren. 2000. "An Approach to Longitudinally Matching Population Survey (CPS)." *Journal of Economic and Social Measurement* 26(1): 31–62.

Marinescu, Ioana. 2017. "The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board." *Journal of Public Economics* 150(C): 14–29.

Modestino, Alicia Sasser, Daniel Shoag, and Joshua Ballance. 2016. "Downskilling: Changes in Employer Skill Requirements Over the Business Cycle." *Labour Economics,* 41: 333–347.

Modestino Alicia Sasser, Daniel Shoag, and Joshua Ballance. Forthcoming. "Upskilling: Do Employers Demand Greater Skill When Workers are Plentiful?" *Review of Economics and Statistics.*

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2017. "Job Changing and the Decline in Long-Distance Migration in the United States." *Demography* 54(2): 631–653.

Mukoyama, Toshihiko, Christina Patterson, and Ayşegül Şahin. 2018. "Job Search Behavior over the Business Cycle." *American Economic Journal: Macroeconomics* 10(1): 190-215.

Osterman, Paul, and Andrew Weaver. 2017. "Skill Demands and Mismatch in U.S. Manufacturing Industries." *Industrial and Labor Relations Review* 70(2): 275–307.

Pew Research Center. 2015. "Searching for Work in the Digital Era." http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/11/PI_2015-11-19-Internet-and-Job-Seeking_FINAL.pdf.

Rothstein, Jesse. 2012. "The Labor Market Four Years into the Crisis: Assessing Structural Explanations*." Industrial and Labor Relations Review* 65(3): 437–500.

Rothwell, Jonathan. 2012. "Education, Job Openings, and Unemployment in Metropolitan America." Metropolitan Policy Program Report. Washington, DC: The Brookings Institution. https://www.brookings.edu/wp-content/uploads/2016/06/29-education-gap-rothwell.pdf.

Rothwell, Jonathan. 2014. "Still Searching: Job Vacancies and STEM Skills." Metropolitan Policy Program Report. Washington, DC: The Brookings Institution. https://www.brookings.edu/wp-content/uploads/2014/07/Job-Vacancies-and-STEM-Skills.pdf.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review* 104 (11): 3529–3564.

Shimer, Robert. 2005. "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy* 113(5): 996–1025.

Shimer, Robert. 2012. "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics* 15(2): 127–148.

Silkroad Technology. 2013. "Recruitment Marketing Effectiveness: Meaningful Metrics Straight From the Source."

Templin, Thomas, and Lesley Hirsch. 2013. "Do Online Job Ads Predict Hiring?" New York: New York City Labor Market Information Services.

Turrell, Arthur, Bradley J. Speigner, Jyldyz Djumalieva, David Copple, and James Thurgood. 2019. "Transforming Naturally Occurring Text Data into Economic Statistics: The Case of Online Job Vacancy Postings." Working Paper No. 25837. Cambridge, MA: National Bureau of Economic Research.

Tüzemen, Didem, and Jonathan L. Willis, 2013. "The Vanishing Middle: Job Polarization and Workers' Response to the Decline in Middle-Skill Jobs." *Economic Review* (Q1): 5–32. https://www.kansascityfed.org/publicat/econrev/pdf/13q1Tuzemen-Willis.pdf.

Vaisey, Stephen. 2006. "Education and Its Discontents: Overqualification in America, 1972– 2002." *Social Forces,* 85(2): 835-864.

Veracierto, Marcelo. 2011. "Worker Flows and Matching Efficiency." *Economic Perspectives* 35(Q4): 147–169.

Zago, Riccardo. 2018. "Job Polarization, Skill Mismatch and the Great Recession." Available at SSRN: http://dx.doi.org/10.2139/ssrn.3004759

Zicklin, Gilbert. 1987. "Numerical Control Machining and the Issue of Deskilling: An Empirical View." *Work and Occupations* 14(3): 452–466.

**Figure 1. Comparing the JOLTS and the BGT Data, January 2010-March 2018**

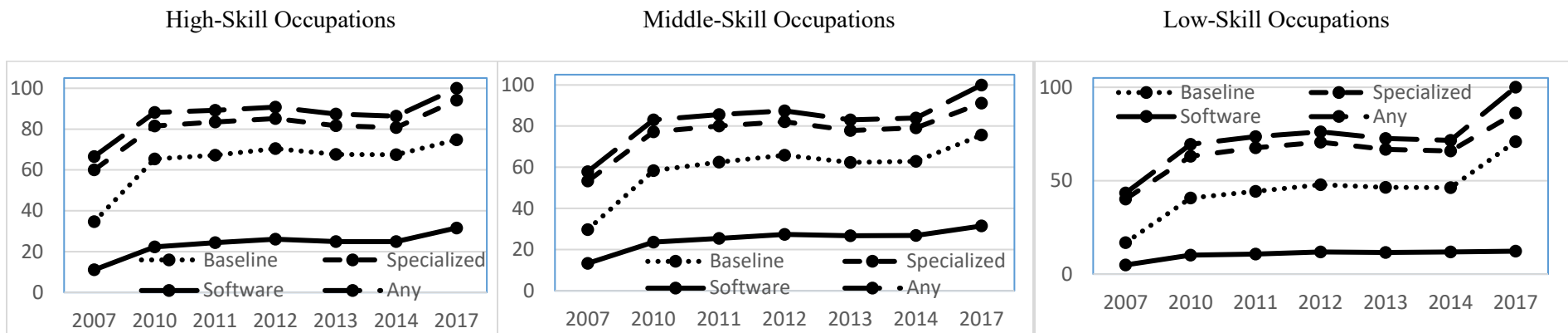(The correlation between the two series is 0.82.)
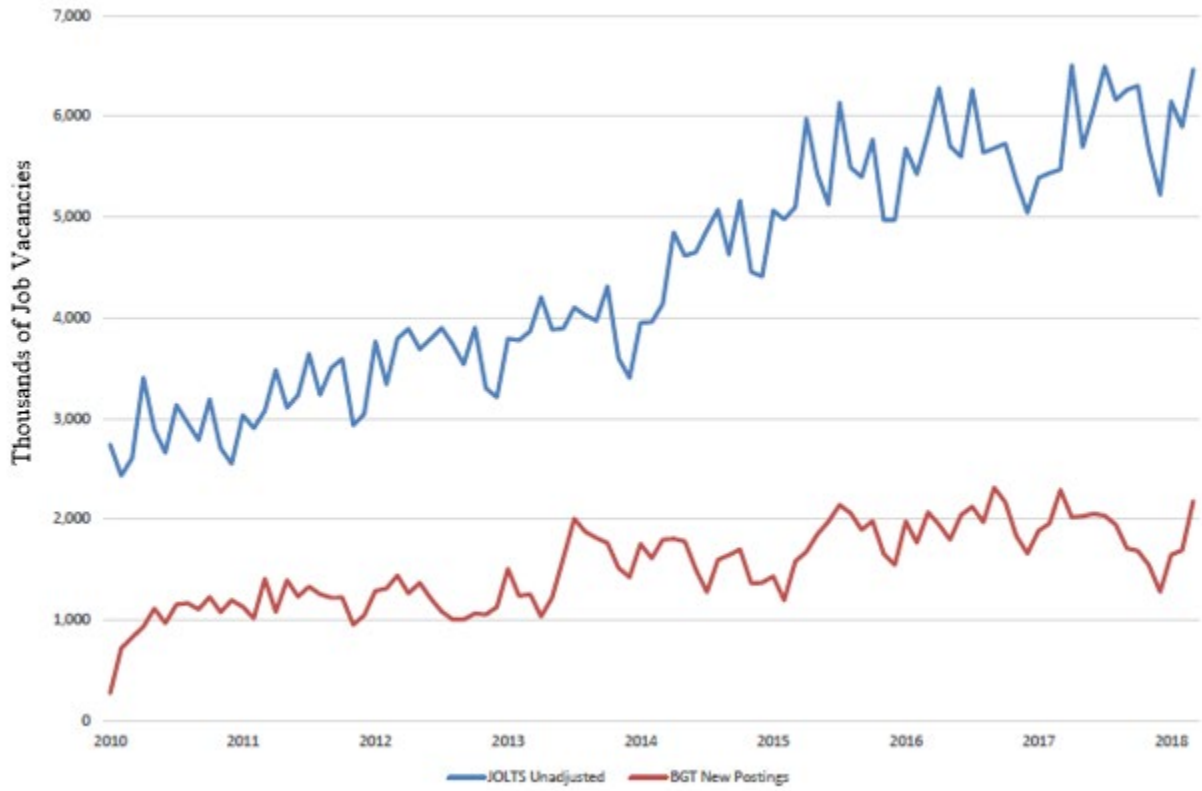
Source: Authors' calculations using data on online job vacancies from Burning Glass Technologies and job vacancies from the Bureau of Labor Statistics Job Openings and Labor Turnover Survey (JOLTS) for 2010-2018.

Figure 2. Trends in Employer Requirements by Skill Sector, 2007 and 2010-2017

A. Share of Postings Requiring Various Levels of Education

High-Skill Occupations   Middle-Skill Occupations   Low-Skill Occupations
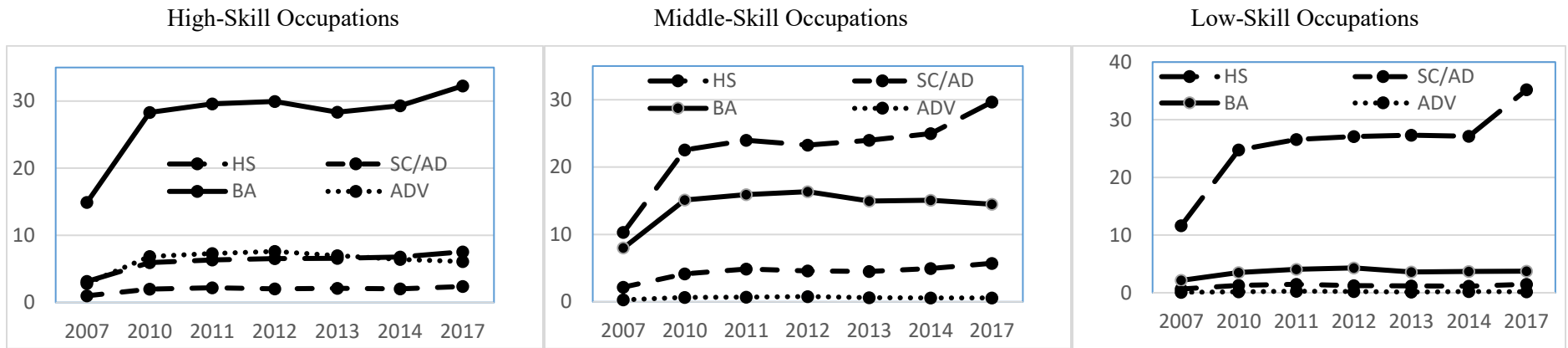
B. Share of Postings Requiring Various Types of Skills
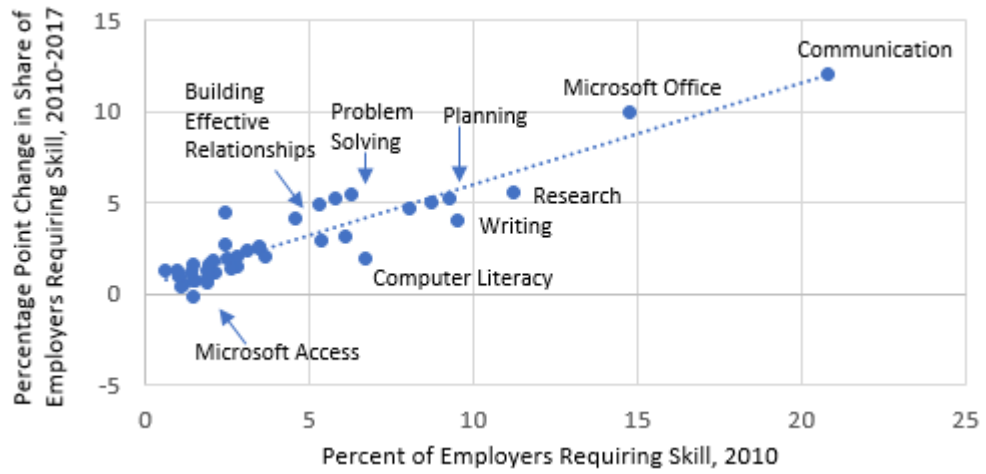
High-Skill Occupations   Middle-Skill Occupations   Low-Skill Occupations

Source: Data on online job vacancies provided by Burning Glass Technologies for 2007 and 2010-2017.

Note: High-skill occupations are defined as those in which at least 40 percent of employees have a four-year college degree or greater according to the 2005-07 American Community Survey. Low-skill occupations are defined as those in which at least 40 percent of employees have a high school education or less according to the 2005-07 American Community Survey. Middle-skill occupations are all other occupations that satisfy neither of these criteria.

37

Figure 3. Initial Level versus Change in Share of Postings Requiring Baseline Skills, 2010 and 2010-17
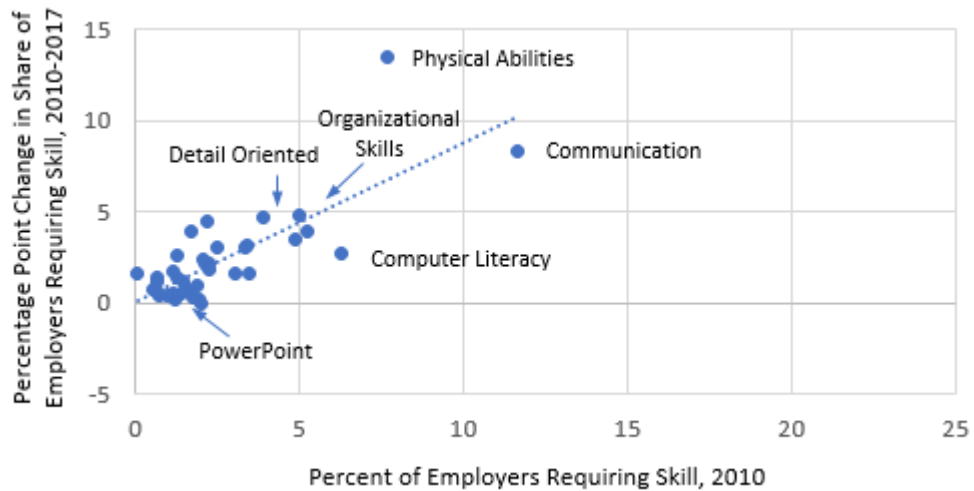
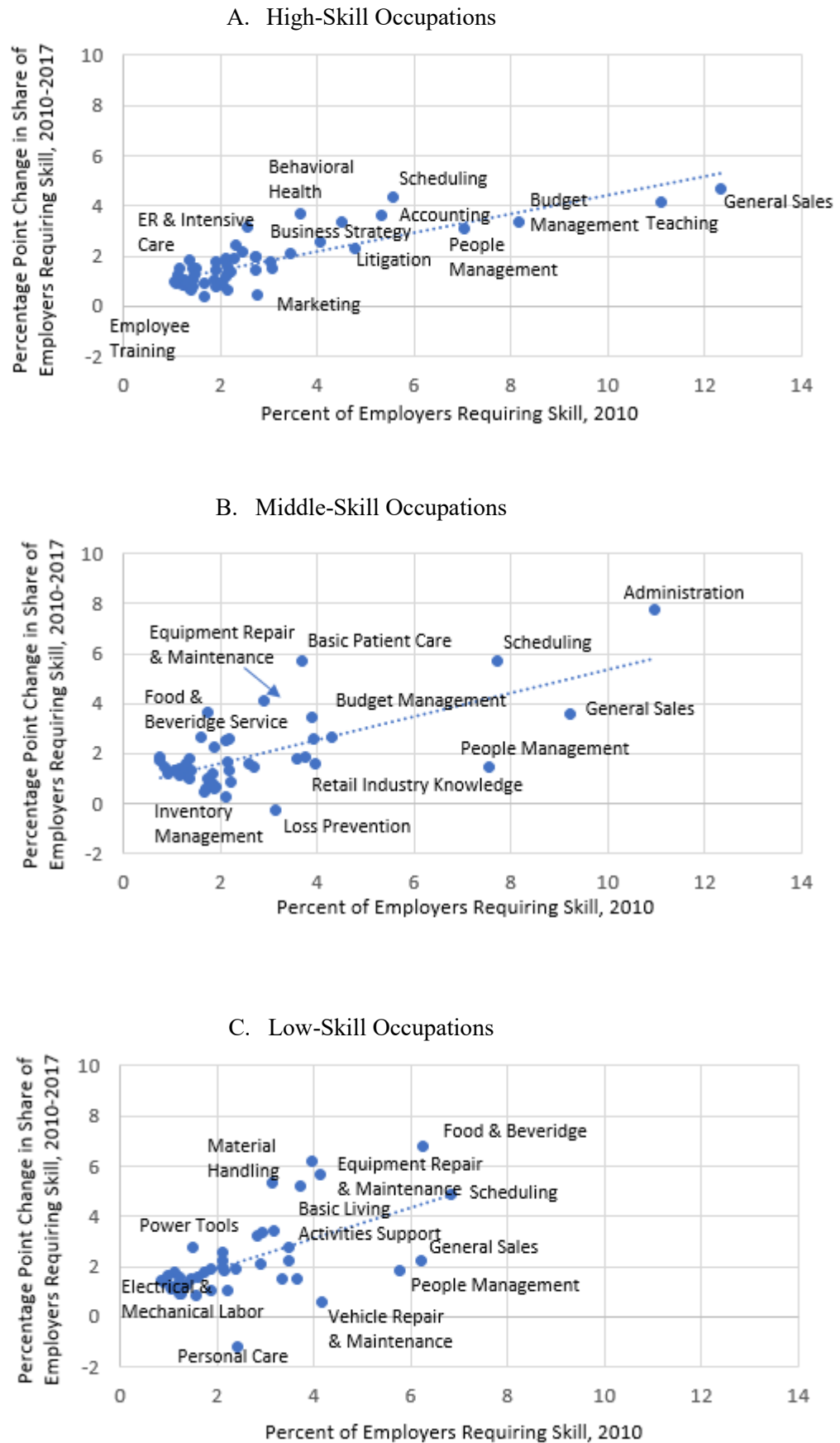A. High-Skill Occupations
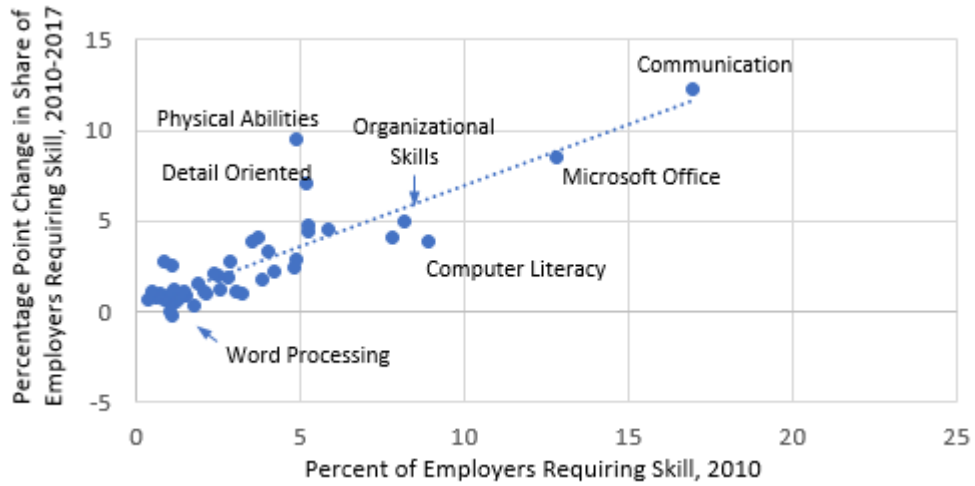
B. Middle-Skill Occupations

C. Low-Skill Occupations

Source: Data on online job vacancies provided by Burning Glass Technologies for 2010 and 2017.

38

Figure 4. Initial versus Change in Share of Postings Requiring Specialized Skills, 2010 and 2017

A. High-Skill Occupations

B. Middle-Skill Occupations

C. Low-Skill Occupations

Source: Authors' calculations using data on online job vacancies from Burning Glass Technologies for 2010 and 2017.

39

Figure 5. Initial versus Change in Share of Postings Requiring Software Skills, 2010 and 2017
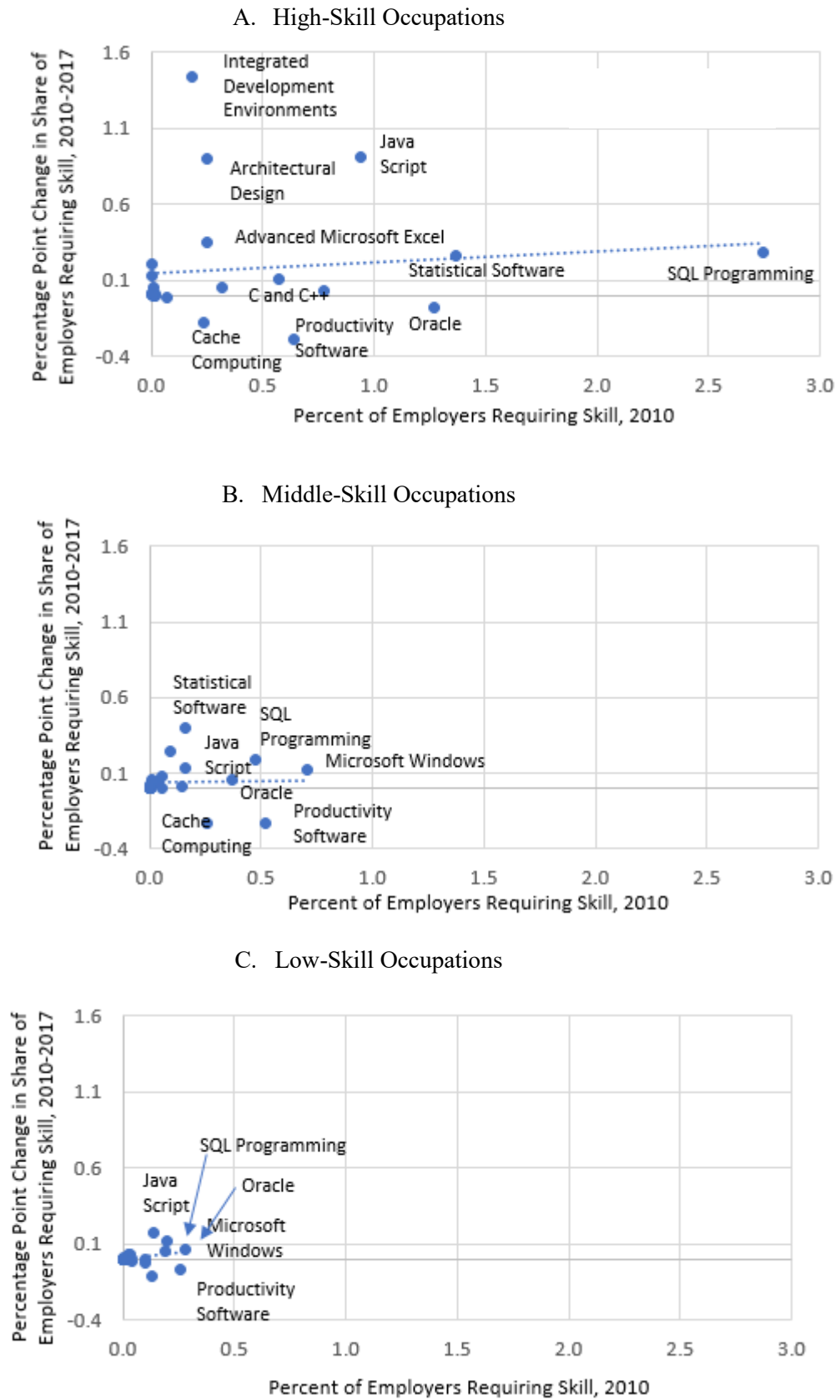
A. High-Skill Occupations

B. Middle-Skill Occupations

C. Low-Skill Occupations

Source: Authors' calculations using data on online job vacancies from Burning Glass Technologies for 2010 and 2017.

40

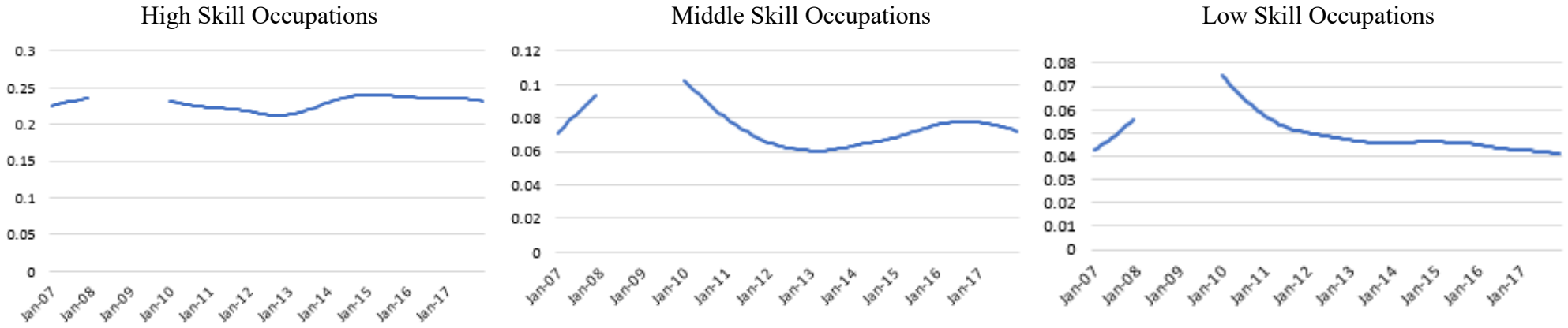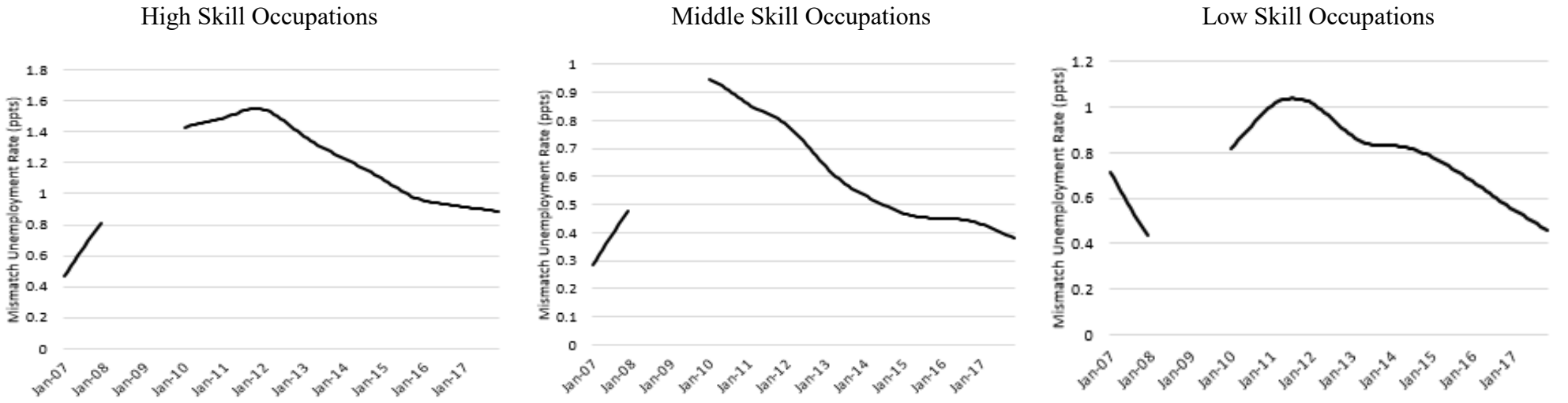**Figure 6. Occupational Mismatch and Unemployment by Skill Groups (2 Digit SOC Level)**

A. Mismatch Index Across 2-Digit SOC Codes

High Skill Occupations	Middle Skill Occupations	Low Skill Occupations

B. Mismatch Unemployment Across 2-Digit SOC Codes

High Skill Occupations	Middle Skill Occupations	Low Skill Occupations

Source: Online job posting data provided by BGT and unemployment and labor force estimates from the CPS for 2007, and 2010 through 2017.
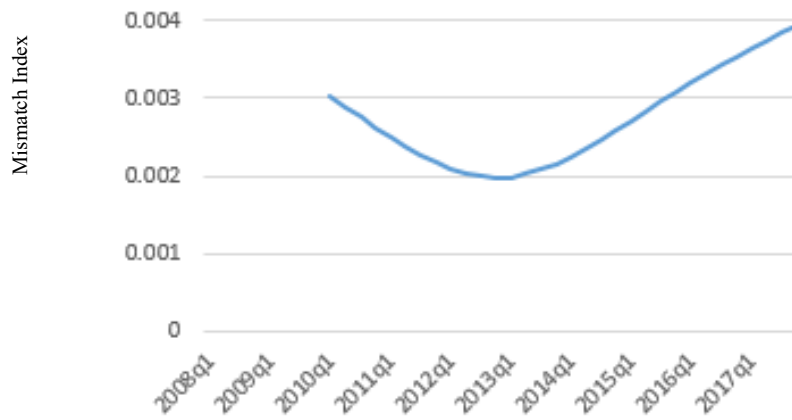Note: The mismatch index and mismatch unemployment rate are calculated following the methodology of Şahin et al. (2014). See the appendix for details.

Figure 7. Mismatch Within 2-Digit High-Skill Occupations that Exhibited Persistent Upskilling

A. SOC 11 – Management

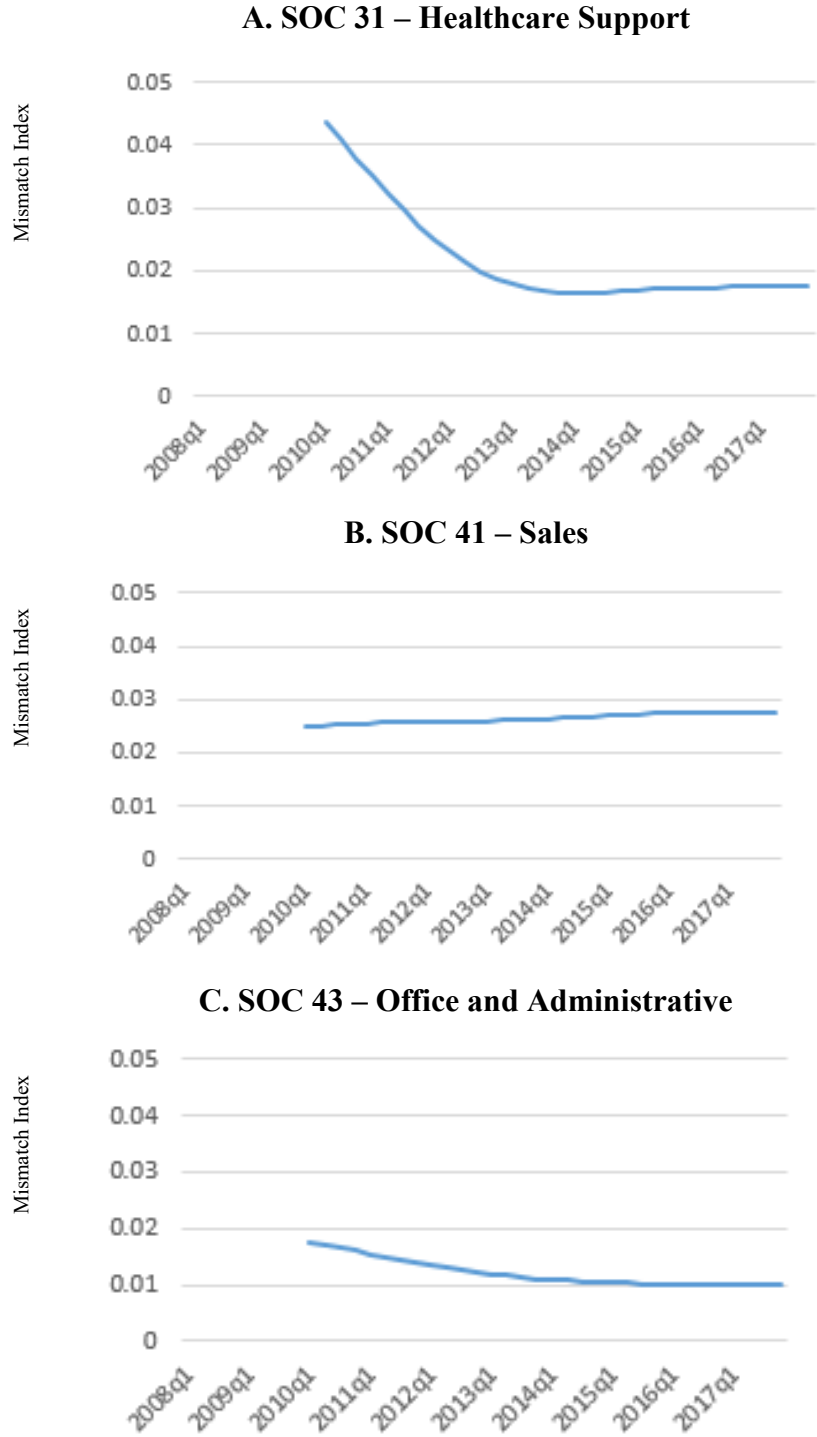**B. SOC 15 – Computer and Mathematical**

C. SOC 13 – Business and Financial

Source: Online job posting data provided by BGT and unemployment and labor force estimates from the CPS for 2010 through 2017.
Note: The analysis was limited to 2-digit SOC codes with at least 1 million job postings in 2010. The mismatch index rate is calculated following the methodology of Şahin et.al. (2014). See the appendix for details.

Figure 8. Mismatch Within 2-Digit Middle-Skill Occupations that Exhibited Temporary/No Upskilling

A. SOC 31 – Healthcare Support

B. SOC 41 – Sales

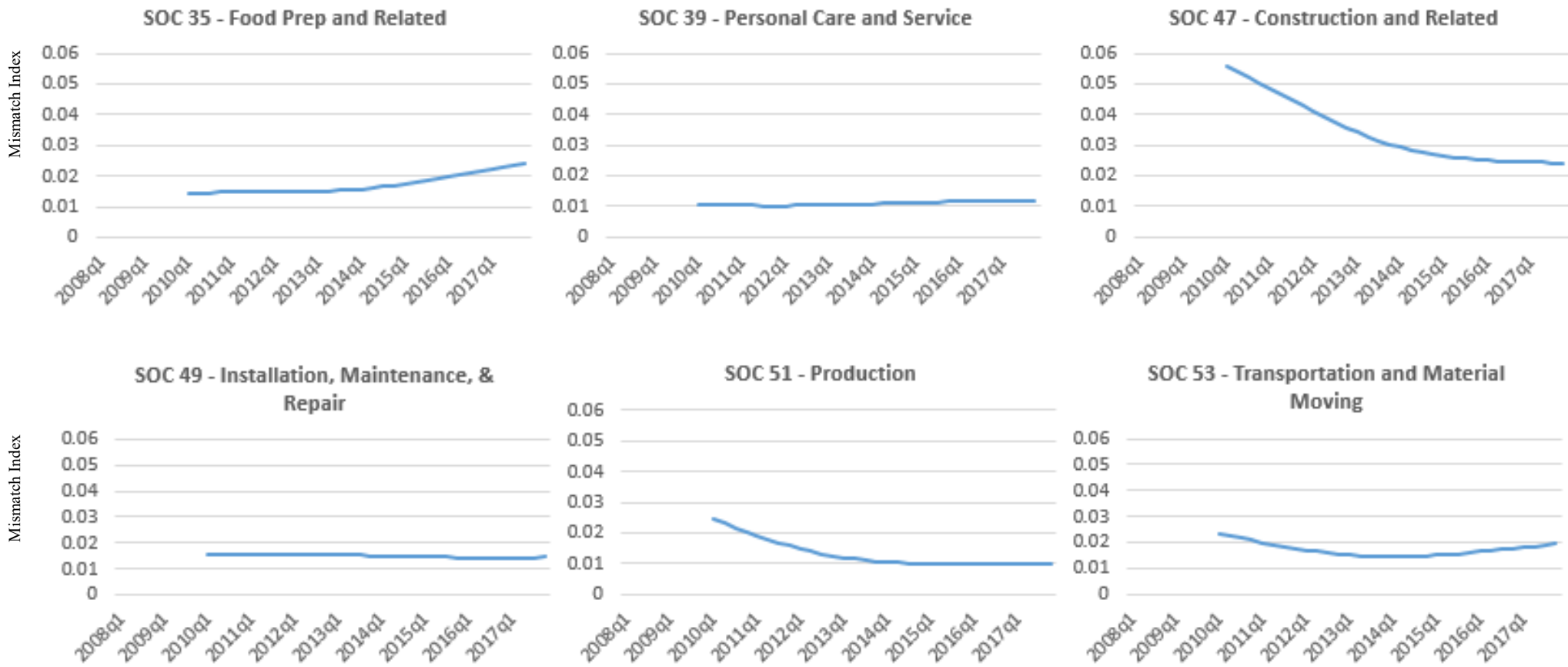**C. SOC 43 – Office and Administrative**

Source: Online job posting data provided by BGT and unemployment and labor force estimates from the CPS for 2010 through 2017.
Note: The analysis was limited to 2-digit SOC codes with at least 1 million job postings in 2010. The mismatch index rate is calculated following the methodology of Şahin et al. (2014). See the appendix for details.

Figure 9. Mismatch Within 2-Digit Low-Skill Occupations that Exhibited Temporary/No Upskilling

Source: Online job posting data provided by BGT and unemployment and labor force estimates from the CPS for 2010 through 2017.
Note: The analysis was limited to 2-digit SOC codes with at least 1 million job postings in 2010. The mismatch index rate is calculated following the methodology of Şahin et al. (2014). See the appendix for details.

Table 1

Heterogeneity in Upskilling Across 3-Digit Occupations by Skill Group, 2007, 2010, and 2017

Panel A. Share of Each Skill Group by Upskilling Type

	Percent of Occupations that Engaged in:		
	Persistent Upskilling	Temporary Upskilling	No Upskilling
All Occupations	39.6	18.7	41.8
Low-Skill Occupations	11.6	18.6	69.8
Middle-Skill Occupations	40.0	25.0	35.0
High-Skill Occupations	82.1	14.3	3.6

Panel B. Distribution of Types of Upskilling by Skill Group

	Percent of Occupations that are:		
	Low-Skill	Mid-Skill	High-Skill
All Occupations	47.3	22.0	30.8
Persistent Upskilling	13.9	22.2	63.9
Temporary Upskilling	47.1	29.4	23.5
No Upskilling	79.0	18.4	2.6

Source: Data on online job vacancies for 2007, 2010, and 2017 are from Burning Glass Technologies.
Note: High-skill occupations are defined as those in which at least 40 percent of employees have a four-year college degree or greater according to the 2005-07 American Community Survey. Low-skill occupations are defined as those in which at least 40 percent of employees have a high school education or less according to the 2005-07 American Community Survey. Middle-skill occupations are all other occupations that satisfy neither of these criteria. An occupation is defined as experiencing upskilling during the Great Recession if there was at least a 5 percentage point increase in the share of postings requiring a four-year college degree or higher during the Great Recession (in our data, between 2007 and 2010). Occupations with less than a 5-percentage point increase in requiring a four-year college degree during the Great Recession are classified as experiencing "no upskilling." We further define an occupation as experiencing "persistent upskilling" if the share of postings requiring a four-year college degree then declined by less than half (2.5 percentage points) during the recovery period (between 2010 and 2017). In contrast, an occupation is defined as experiencing "temporary upskilling" if the share of postings requiring a four-year college degree were reversed by a more than half (2.5 percentage-points) during the recovery period. Shares are weighted by the occupation's share of total employment as of 2006.

Table 2. Difference-in-Difference Analysis of Skill Requirements: Permanent v. Temporary Upskilling Occupations, 2007, 2010, and 2017

	Change in Skill Requirements					
	Recession: 2007 to 2010			Recovery 2010 to 2017		
	Share Requiring a BA					
	2007	2010	Diff	2010	2017	Diff
Persistent Upskillers	19.49	37.29	17.80 ***	37.29	41.11	3.82
Temporary Upskillers	13.87	27.43	13.56 ***	27.43	14.88	-12.55 ***
Diff	-5.61	-9.86		-9.86	-26.23	
Diff-in-Diff			-4.24			-16.37 ***
	Share Requiring Baseline Skills					
	2007	2010	Diff	2010	2017	Diff
Persistent Upskillers	38.56	68.61	30.05 ***	68.61	79.98	11.38 ***
Temporary Upskillers	33.74	65.12	31.38 ***	65.12	76.03	10.91
Diff	-4.82	-3.49		-3.49	-3.96	
Diff-in-Diff			1.33			-0.47
	Share Requiring Specialized Skills					
	2007	2010	Diff	2010	2017	Diff
Persistent Upskillers	58.07	81.09	23.03 ***	81.09	93.39	12.29 ***
Temporary Upskillers	47.46	73.56	26.11 ***	73.56	89.30	15.73
Diff	-10.61	-7.53		-7.53	-4.09	
Diff-in-Diff			3.08			3.44
	Share Requiring Software Skills					
	2007	2010	Diff	2010	2017	Diff
Persistent Upskillers	13.48	24.97	11.49 ***	24.97	33.85	8.88 **
Temporary Upskillers	9.29	20.23	10.93 ***	20.23	20.75	0.52
Diff	-4.19	-4.74		-4.74	-13.11	
Diff-in-Diff			-0.56			-8.37 ***
	Share Requiring Any Skills					
	2007	2010	Diff	2010	2017	Diff
Persistent Upskillers	65.62	88.11	22.49 ***	88.11	100.00	11.89 **
Temporary Upskillers	51.92	80.83	28.92 ***	80.83	100.00	19.17 **
Diff	-13.71	-7.28		-7.28	0.00	
Diff-in-Diff			6.43 *			7.28 **

Source: Data on online job vacancies provided by Burning Glass Technologies for 2007, 2010, and 2017.
Notes: An occupation is defined as having upskilled during the Great Recession if there was at least a 5 percentage point increase in the share of postings requiring a four-year college degree or higher during the Great Recession (between 2007 and 2010). We further define an occupation as experiencing "persistent upskilling" if the share of postings requiring a four-year college degree then declined by less than half (2.5 percentage points) during the recovery period (between 2010 and 2017). In contrast, an occupation is defined as experiencing "temporary upskilling" if the share of postings requiring a four-year college degree were reversed by a more than half (2.5 percentage-points) during the recovery period.

Table 3. Change in Mismatch Unemployment by Skill Sector, 2010 – 15

	Mismatch Unemployment, 2010 PPT	Mismatch Unemployment, 2010 Percent	Mismatch Unemployment, 2015 PPT	Mismatch Unemployment, 2015 Percent
Low-Skill Occupations	0.87	7	0.46	6
Middle-Skill Occupations	0.94	14	0.41	12
High-Skill Occupations	1.48	31	0.82	47

Source: Online job vacancy data provided by Burning Glass Technologies, unemployment and labor force estimates from the CPS.
Note: The abbreviation "PPT" means that the amount of mismatch unemployment is expressed in raw percentage points, and "Percent" means that mismatch unemployment is expressed as a percentage of the counterfactual unemployment rate that would occur in the absence of any labor market mismatch. The mismatch unemployment rate is calculated using the methodology of Sahin et al. (2014). See the appendix for details.

ONLINE APPENDIX

I. DATA SOURCES

A. Online Job Vacancy Data from Burning Glass Technologies

1. General Description

The main source of the data used in this paper is collected by Burning Glass Technologies (BGT), one of the leading vendors of online job posting data.[23] On a daily basis, BGT collects detailed information on over 7 million current online job openings gathered from over 40,000 sources, including job boards, online news media, government agencies, and employer websites. BGT's collection process provides a robust representation of online hiring activity, including job listings posted by small employers.

The data are collected via a web crawling technique that uses computer programs called "spiders" to browse online job boards and other web sites and systematically parse the text in each job listing into usable data elements. The "spidering" process follows a fixed schedule of browsing a pre-determined basket of websites that are carefully monitored and updated to include the most current and complete set of online postings. BGT has developed algorithms that identify a series of identically parsed variables across job ads, such as location, employer, and job title, to eliminate the duplicate counting of a single ad—for example, one that is posted on an employer website as well as on a large job board. To avoid abrupt and possibly erroneous fluctuations in vacancies over time, BGT places more weight on large job boards than individual employer sites, as the latter are updated less frequently.

BGT mines over 70 job characteristics from the text of job postings, including the employer's name, location (down to the level of the MSA), job title, occupation (down to the 6-digit Standard Occupational Classification code, or SOC), years of experience requested, and level of education required or preferred by the employer. As such, the data are unique in allowing a geographic analysis of occupation-level labor demand by education level and experience level. Note that the BGT data do not contain any information on the duration of the vacancy, how many applications a vacancy received, nor whether a vacancy was ever filled. Over the nine years of data

---

[23] See http://www.burning-glass.com/realtime/ for more details.

that we examine (2007 and 2010–2017), the BGT data represent a total of roughly 159 million vacancies.

We make use of two different products from BGT: a dataset of individual vacancies that BGT pooled for each entire year without duplicating any specific posting, and a normalized dataset that is pegged to the number of vacancies found in the JOLTS (Job Openings and Labor Turnover Survey), both in the aggregate and by industry. We also have access to BGT's Labor/Insight analytical tool, which enables us to validate many of the important components of the pooled data, including timeframes, de-duplication, and aggregation.

2. Representativeness and Consistency of the BGT Data

In using the BGT online vacancies data, we acknowledge the concerns that (1) the data may not be representative of job vacancies in general because the data only consist of online vacancies, and (2) the data may not be representative of the universe of online job postings. In addition, the nature and representativeness of the BGT data may have changed within the time period we observe, which consists of the year 2007 and the years 2010–2017. Regarding the internal consistency of the data over time, although BGT regularly improves its data-collection algorithms, any improvements are applied retroactively to the complete historical database of postings. Therefore, we do not face any issues stemming from changes over time in how the data were collected and processed by BGT.

Concerning representativeness, as constructed, the BGT data fail to capture job vacancies that are posted in offline sources such as print newspapers, television or radio ads, "Help Wanted" signs, or via word-of-mouth. In addition, a single online ad might represent multiple openings with the same job description, thus introducing a further source of undercounting. Therefore, the omissions mean that the BGT data will understate the total number of job openings in the United States, and the shortfall may differ by industry or occupation according to the tendency of different employers to post vacancies online. We address these concerns by comparing the BGT data to the JOLTS data and data obtained from the Minnesota Job Vacancies Survey.[24]

Figure 1, appearing in the set of figures from the main text, shows that the BGT data exhibit a high degree of correlation over time with the JOLTS vacancy series (0.82), despite the fact that,

---

[24] The Minnesota Job Vacancy Survey is one of twelve state job vacancy surveys conducted in the United States, collected biannually since 2001.

as expected, the level of BGT vacancies is consistently lower. Furthermore, movements in the number of postings over time between 2007–2010 and 2010–2012 were roughly similar between the full BGT dataset and the JOLTS, and were nearly identical between the Minnesota Jobs Vacancy Survey and the BGT data for Minnesota, as shown in Table A1 below.

Given the inherent shortfall in the level of BGT vacancies relative to the JOLTS, and to obtain an accurate number of vacancies over time, BGT creates a normalized (reweighted) dataset that exactly matches the monthly JOLTS vacancies by industry. We use this second version of the data when calculating the number of vacancies by occupation on the demand side to construct our mismatch indices by skill sector.

Although we do not classify vacancies by industry, it is important to examine how representative the BGT data on vacancies by industry is over time, as this could affect how accurate the BGT vacancies series by occupation is over time. Figure A1 plots the distribution of BGT postings for 2007 across major industry groups, sorted from largest to smallest, as well as the distribution of job vacancies in the JOLTS. Despite the differences in the sampling methods used by these two sources, the industry distributions are quite similar—the BGT data deviate by less than 10 percentage points from the JOLTS data across all major categories. The largest deviations show that the BGT data tend to over-represent industries such as finance and insurance, healthcare and social assistance, and educational services, and to underrepresent food services, public administration/government, and construction. Again, most of these differences are small in magnitude.

However, we are also concerned with whether the representativeness across industries has changed over time. Figure A2 compares the industry distribution of the BGT data to that of the JOLTS data in both 2007 and 2012. On the x-axis we plot the deviation of the BGT industry share in 2007 from that industry's share of job vacancies in the JOLTS data during the same year. In 2007, for example, the finance and insurance industry, shown on the far right, is over-represented by roughly 7 percentage points in the BGT data as compared to JOLTS, and accommodation and food services, on the far left, is under-represented by roughly 7 percentage points in the BGT data when compared to the JOLTS data. On the y-axis we plot the same deviation for 2012. We also plot the 45-degree line as a benchmark: if the representativeness of the BGT data, relative to the JOLTS, remained constant over time, all markers should line up on the 45-degree line.

Figure A2 shows that changes in the data's representativeness over this time period are very small, with most of the markers lining up closely to the 45-degree line. To the extent that any changes did occur in the BGT data, there is a tendency for these data to more closely approximate the JOLTS data. In 2012, for example, finance and insurance was overrepresented by only 4 percentage points in the BGT data, while accommodation and food services was underrepresented by 4.5 percentage points in the BGT data.

The primary advantage of using the BGT data is that they allow us to categorize vacancies by occupation at a detailed level, which is not possible to do with the JOLTS data. We can assess how representative the BGT postings are by occupation if we compare the BGT data for Minnesota with postings from the Minnesota Job Vacancy Survey (MN JVS), which also includes the occupation for each vacancy. Figure A3 plots the distribution of postings across broad occupational groupings, sorted from largest to smallest, for each of these data sources. As with the JOLTS comparison, the occupational distributions are quite similar despite the differences in the sampling methodology. Compared to the MN JVS, the deviation in the BGT data is, again, less than 10 percentage points across all major occupational categories. The largest deviations show that the BGT data tend to overrepresent higher-skilled occupations, such as computer and mathematical professions (+9 percentage points), management (+7 percentage points), and business and financial (+6 percentage points), while underrepresenting lower-skilled occupations such as food preparation and serving (–7 percentage points), healthcare support (–5 percentage points), and transportation and material moving (–4 percentage points). Again, most differences are small in magnitude and our findings are similar to those of other researchers.[25]

Again, we are concerned with whether the representativeness of job openings across occupations has changed over time. Figure A4 compares the occupational distribution of the BGT data to that of the MN JVS in both 2007 and 2012, similar to the comparison we did earlier with the JOLTS data and the BGT data. As with the industry comparisons, most of the markers lie close to the 45-degree line, indicating that any differences in the representation of occupations between the BGT data and the MN JVS data were fairly stable between 2007 and 2012.

---

[25] Rothwell (2014) also compares the occupational distributions from an extracted sample of the BGT data to those distributions obtained from state vacancy surveys for selected metropolitan areas for which data are available. He finds that computer, management, and business occupations are overrepresented relative to the state vacancy surveys, while healthcare support, transportation, maintenance, sales, and food service jobs are underrepresented.

Concerning how representative the BGT data are with respect to just the universe of online postings, BGT's coverage depends on the number of websites it scrapes, as well as on the proprietary algorithms it uses to recognize and eliminate duplicate job postings and to accurately code information (such as occupation) in the postings. Carnevale, Jayasundera, and Repnikov (2014) show that the occupation-industry composition of the BGT data are similar to that of the Conference Board's Help Wanted Online Index (HWOL) of vacancies. Moreover, comparing a sample of job postings in the BGT database to the actual text of the same postings, the same authors concluded that the BGT coding for occupation, education, and experience was accurate for 80 percent of the listings that were audited. The accuracy of the data that we employ in this paper is likely to be even greater, because BGT has since made improvements to its coding algorithms and retroactively applied these fixes to the earlier data.

In order to further assess the similarity between the BGT vacancies data and the HWOL vacancies data, we compared our own estimates of aggregate occupational mismatch (by major 2-digit SOC code) using the BGT vacancies data to the previous estimates of occupational mismatch (conducted by Şahin et al. 2014) that employed the HWOL vacancies data. We found that for the overlapping time periods (2007 and 2010–2012), the respective estimates (based on comparing either the unadjusted or adjusted mismatch indexes) lie within a similar range and exhibit qualitatively similar movements. Our results are available on request.

3. Remaining Questions about the BGT Data

We acknowledge that there are still some remaining questions regarding how to think about online postings. For example, the BGT data do not indicate how many of the observed vacancies were ever filled. Nonetheless, a few recent studies suggest that online vacancies are highly salient sources of information for both firms and jobseekers, and that online job postings are indicative of actual hiring activity. Data collected by SilkRoad Technology, a global provider of social talent management solutions, indicates that online job postings have become the primary source of advertising open positions for firms.[26] In addition, a 2015 survey by the Pew Research Center found that online employment resources now rival personal and professional networks as a top

---

[26] SilkRoad Technology. 2013. "Recruitment Marketing Effectiveness: Meaningful Metrics Straight From the Source."

source of job information for Americans who are looking for work.[27] Concerning the relationship between online vacancies and hiring activity, Templin and Hirsch (2013) find that, at least in New York City during the Great Recession and through 2010, the number of online job ads was a strong predictor of actual hiring activity one quarter later (Templin and Hirsch 2013). When specifically examining the BGT data, Hershbein and Kahn (2018) find that the occupational composition of BGT online postings agrees strongly with the distribution of new hires in the CPS between 2007 and 2015.

Other researchers have also provided suggestive evidence that the occupational distributions in the BGT data reflect that of new hires, implying that there is a reasonable connection between online job postings and hiring activity. For example, Hershbein and Kahn (2018) validate this aspect of the BGT data by comparing the occupational distribution of jobs postings contained in the BGT to the new jobs appearing in the Current Population Survey over time. They find that computer and mathematical occupations, management occupations, and architecture and engineering occupations appear to have become less overrepresented, while healthcare and business and finance look fairly unchanged. In contrast, administrative support, food, transportation, and production occupations have become slightly less underrepresented. For most of these occupations, though, the differences they cite are "quite small."

---

[27]Pew Research Center. 2015. "Searching for Work in the Digital Era. " http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/11/PI_2015-11-19-Internet-and-Job-Seeking_FINAL.pdf

Figure A1. Industry Distribution of Job Postings: BGT versus JOLTS, 2007
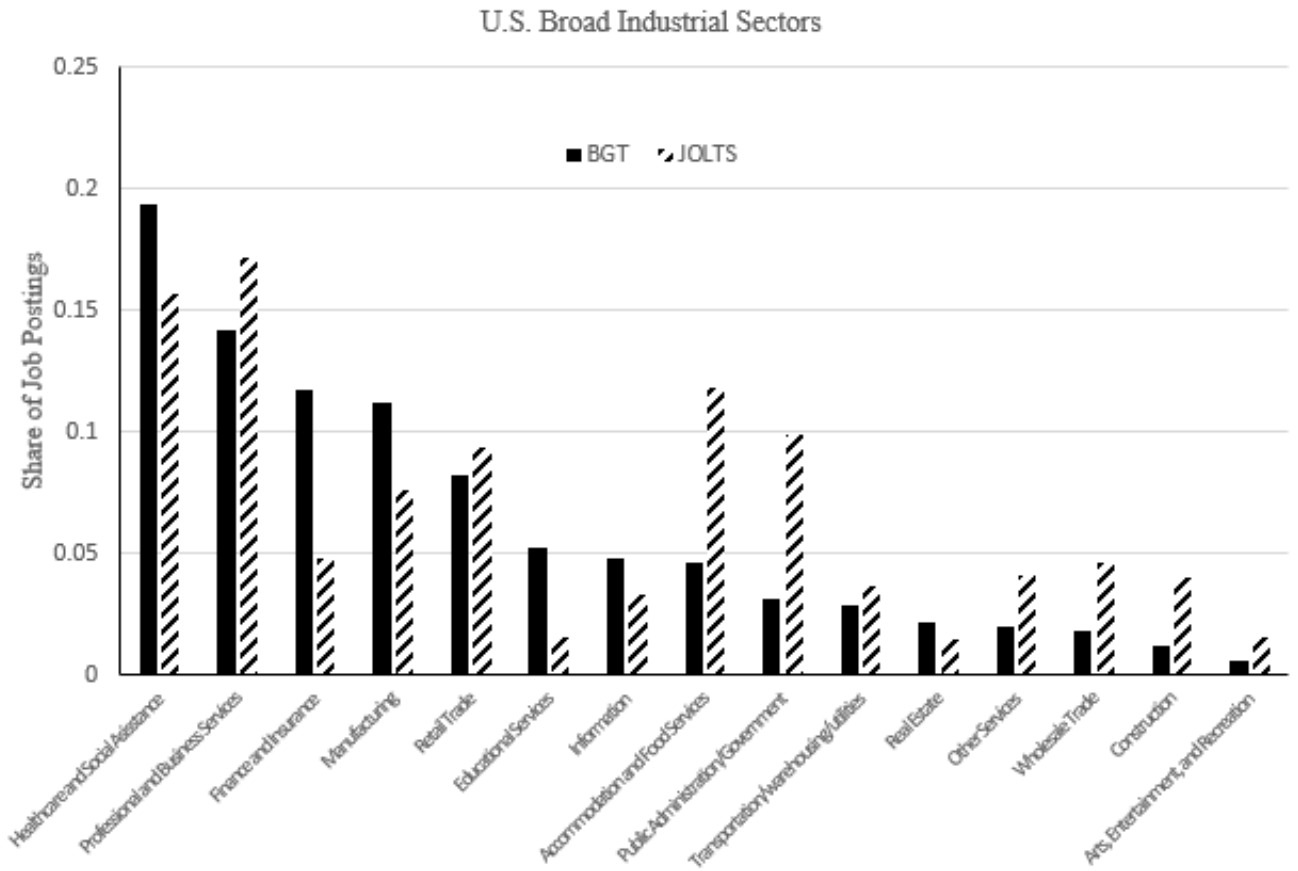
U.S. Broad Industrial Sectors

Source: Authors' calculations using data from Burning Glass Technologies and the BLS Job Openings and Labor Turnover Survey, 2007
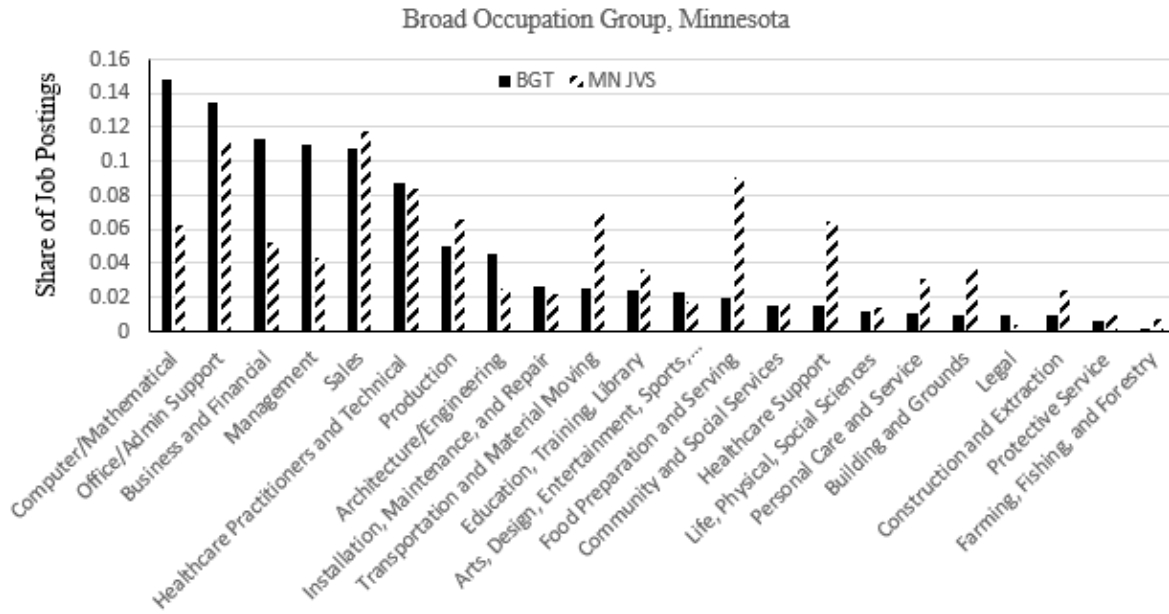
Figure A2. Industry Distribution of Job Postings: BGT versus JOLTS, 2007 and 2012
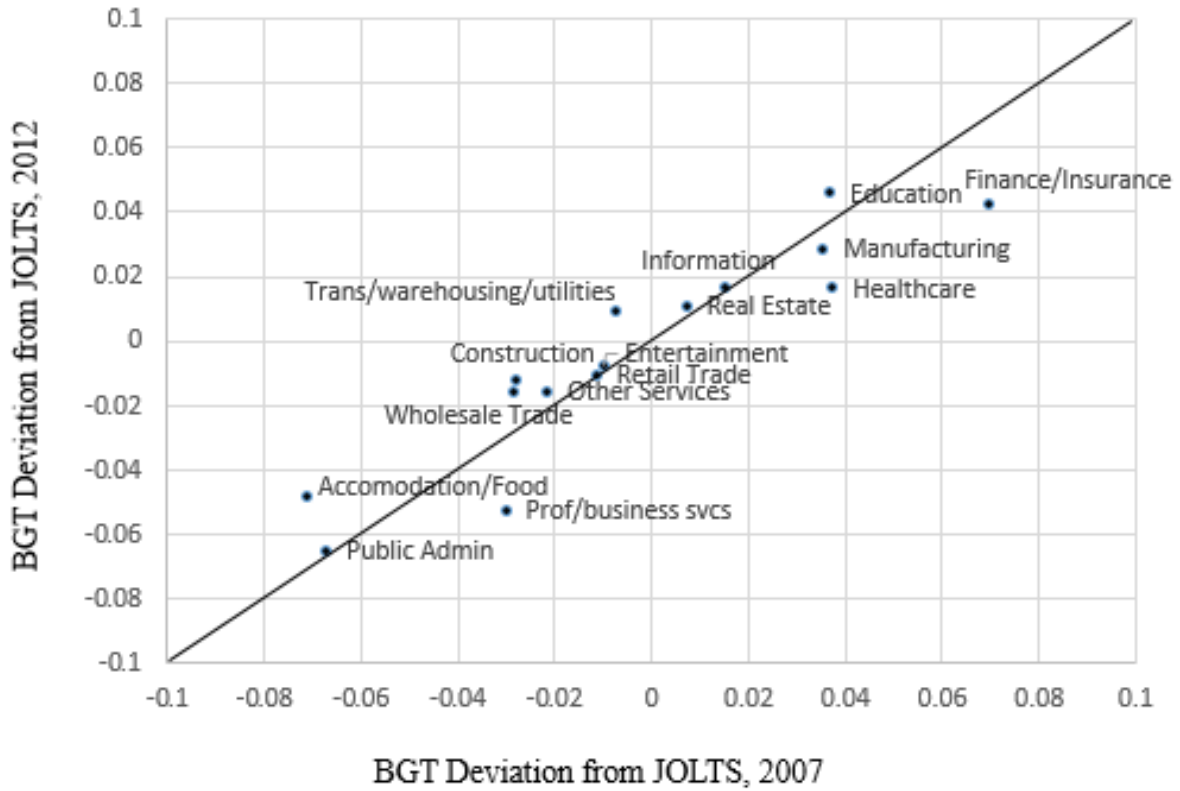
Source: Author's calculations using data from Burning Glass Technologies and the BLS Job Openings and Labor Turnover Survey, 2007 and 2012.
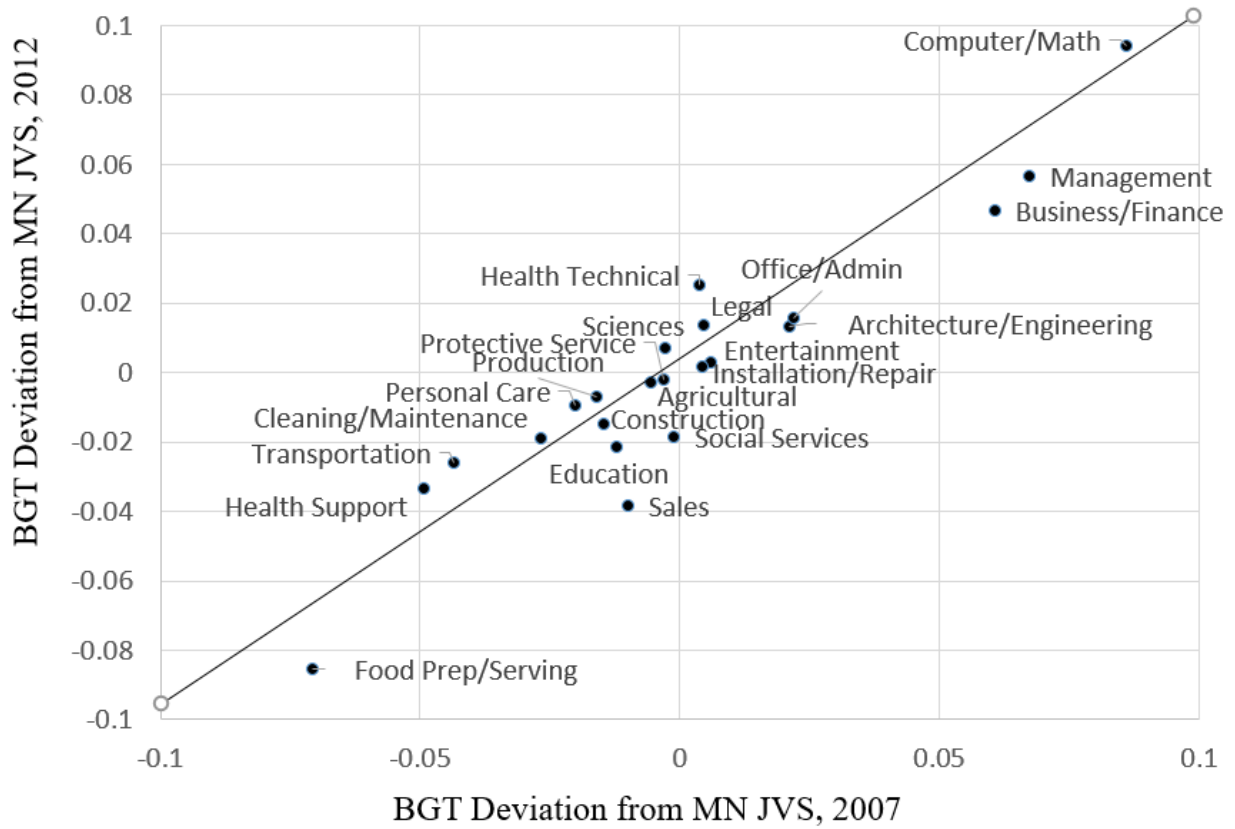Note: The x-axis is the BGT job vacancy share in an occupation in 2007 minus the JOLTS job vacancy share in the same industry in 2007. The y-axis is these same differences for 2012. As a benchmark, the 45-degree line (black dash) indicates industries where representation in BGT, relative to the JOLTS, did not change from 2007.

Figure A3. Occupation Distribution of Job Postings: BGT versus MN JVS, 2007

Source: Authors' calculations using data from Burning Glass Technologies and the Minnesota Job Vacancy Survey, 2007.

Figure A4. Occupation Distribution of Job Postings: BGT versus MN JVS, 2007 and 2012

Source: Author's calculations using data from Burning Glass Technologies and the Minnesota Job Vacancy Survey, 2007 and 2012.
Note: The x-axis is the BGT job vacancy share in an occupation in 2007 minus the MN JVS job vacancy share in the same occupation in 2007. The y-axis is these same differences for 2012. As a benchmark, the 45-degree line (black dash) indicates occupations where representation in BGT, relative to the MN JVS, did not change from 2007.

Table A1. Comparison of Overall Trend in Job Openings versus Online Postings: 2007, 2010, and 2012

		Level			Percent Change	
		2007	2010	2012	2007 to 2010	2010 to 2012
Total United States						
	JOLTS	4477.5	2852.6	3671.9	-36.3%	28.7%
	BGT	1060.4	786.7	1057.6	-25.8%	34.4%
Minnesota						
	MN	56.6	37.6	60.9	-33.5%	62.0%
	BGT	56.0	36.5	58.2	-34.9%	59.6%

Source: Author's calculations using data from the BLS Job Openings and Labor Turnover Survey (JOLTS), Burning Glass Technologies (BGT), and the Minnesota Job Vacancy Survey (MN JVS).
Note: The level of job openings (JOLTS and Minnesota JVS) and the level of job postings (BGT) are reported in thousands. Comparisons to the JOLTS data are for average monthly job postings over the year.

B. Current Population Survey

1. General Description

The Current Population Survey (CPS) is a monthly survey of about 60,000 US households conducted by the United States Census Bureau for the Bureau of Labor Statistics (BLS).[28] We use the IPUMS-CPS microdata for the Current Population Survey to generate estimates of the number of unemployed individuals by either minor occupation or detailed occupation, as described below, and these estimates enter into our calculations of the various mismatch indexes. Separately, in the context of conducting robustness checks we use the IPUMS-CPS microdata to first identify new hires and continuing employees, and then to describe trends from 2005 through 2017 in the relative education level of new hires versus continuing employees for each skill sector.

2. Identifying Unemployed Workers by Occupation

In the context of calculating the mismatch indexes by skill sector, we quantify the number of unemployed jobseekers by detailed occupation (6-digit Standard Occupation Classification, or SOC) per month and, as described in Section II below, we aggregate these numbers up to the level of the major occupation (2-digit SOC) by skill sector. To obtain the number of unemployed workers by detailed occupation, we use the monthly IPUMS-CPS, which indicates whether an individual was unemployed, as opposed to being either employed or not in the labor force. For those individuals counted as unemployed, the CPS indicates the detailed occupation code for the worker's most recent job, which is assumed to be the occupation in which the worker is seeking a new job.[29] To match the detailed occupation codes in the CPS to the 6-digit SOCs, we use the crosswalk provided in Şahin et al. (2014).

Separately, in order to calculate the mismatch indexes across minor occupations (3-digit SOC codes) within major occupations, we use the CPS data to quantify the number of unemployed workers per minor occupation in a given month. However, due to the problem of missing

---

[28] See https://www.census.gov/programs-surveys/cps.html for more information.

[29] Occupational information for unemployed workers in the CPS is missing in a non-trivial fraction of cases, and we drop all such observations when calculating the mismatch indexes. This approach may introduce measurement error in our mismatch calculations, and as noted below we take some steps to mitigate the problem. Şahin et al. (2014) also show that estimates of mismatch across 2-digit occupations are not significantly different when they drop unemployed individuals with missing occupation data as opposed to imputing occupations to such individuals.

information on the worker's previous occupation (see footnote 7 above), the unemployment numbers for some 3-digit SOC codes exhibit implausible fluctuations across months, so we use the quarterly average unemployment numbers by 3-digit SOC code instead. Cases in which we observe that there are no unemployed individuals in a given 3-digit SOC code for a given month are treated as missing values, so are dropped from the calculation of the relevant quarter's average. In the case of 6-digit SOC codes for a particular month, we also observe cases in which there are no unemployed persons in a given occupation, but this issue is mitigated by the fact that we aggregate the unemployment numbers back up to the 2-digit SOC level by skill sector. See Section II below for more information.

3. Classifying Occupations as Low-Skill, Middle-Skill, or High-Skill

We use the CPS to quantify the number of unemployed individuals by the detailed 6-digit SOC code in order to calculate our mismatch indexes by skill sector. To aggregate these numbers up to the level of the 2-digit SOC code by skill sector, we first assign a skill category to each 6-digit SOC code according to the predominant level of educational attainment for incumbent workers in that occupation using data from the American Community Survey (ACS) as of 2005–2007. In particular, "high-skill" 6-digit SOC codes are defined as those in which at least 40 percent of the individuals employed in the occupation (as of the pooled 2005–2007 ACS) had at least a four-year college degree. "Low-skill" occupations are those in which at least 40 percent of employed workers (as of the 2005–2007 ACS) had a high school diploma or less, and "middle-skill" occupations are those that meet neither of these two educational criteria, indicating a more mixed education profile among employees. These definitions, which we hold fixed throughout our analysis, result in no ambiguous assignments—each occupation meets exactly one of the definitions. For a description of how we aggregate the unemployment counts from the 6-digit SOC level up to the level of the 2-digit occupational code by skill sector, see Section II below.

4. Calculating the Average Education Level of New Hires Relative to the Average Education Level of Continuing Employees by Skill Sector

In order to calculate the relative education level of new hires versus continuing workers by skill sector, we first use the CPS to identify new hires as well as continuing employees. We identify new hires and continuing employees using the methods of longitudinal matching established by

Madrian and Lefgren (2000) and Hershbein (2017). An individual is deemed to be a new hire in a given month if, since the previous month, the individual moved either (1) from being not in the labor force to employed, (2) from unemployed to employed, or (3) from one employer to another. We do not count as new hires those individuals who changed occupations within the same employer, as such changes can show up for spurious reasons and including them could exaggerate the count of new hires. An individual is considered to be a continuing employee if he or she was employed in both the current month and the previous month and did not experience a change of employer between these two months.

For each new hire or continuing employee, using the CPS we observe the individual's detailed occupation and education level, provided such information is not missing. Based on the detailed occupation, the observation is assigned to the appropriate skill sector using the procedure described below in Section II. We assign a discrete value to the education level of each observation—1 for high school or less, 2 for some college or an associate's degree, 3 for a four-year college degree, and 4 for an advanced degree. Then, separately by skill sector and by year for the period from 2005 through 2017, we calculate the (weighted) average education level among new hires and the (weighted) average education level among continuing employees, using the longitudinal weights furnished in the CPS. Finally, we divide the average education level of new hires (for a given skill sector and year) by the average education level of continuing employees (for the same sector and year) to obtain an estimate of the relative average education level of new hires versus continuing employees in each year. The results of these calculations are discussed in Section III.A below.

II. Estimation of Mismatch Indices and Mismatch Unemployment

In order to compute the value of the mismatch index, either across major occupations within a specific skill sector or across minor occupations that fall within a major occupation, we use equation (3), shown in the main text. For the occupation-specific matching efficiency para-meters we use the values given in the online appendix of Şahin et al. (2014). We estimate the remaining model parameters—including the vacancy elasticity of hires, aggregate matching efficiency, and occupation-level productivity parameters—using the methods and data described in the online appendix to Şahin et al. (2014). To calculate mismatch unemployment, we follow the equations and methods described in Şahin et al. (2014), conducting each calculation for the
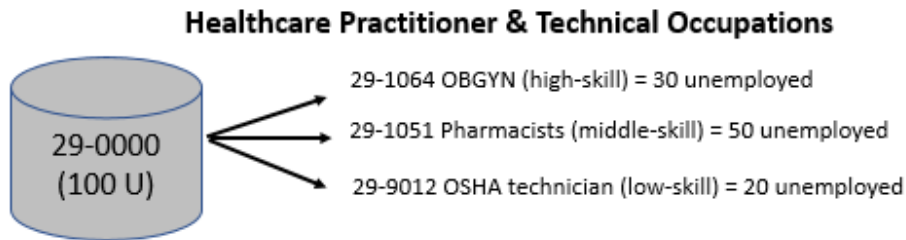
appropriate skill sector or major occupation, depending on the exercise. In calculating mismatch unemployment, we must estimate the counterfactual unemployment levels that would occur in the absence of mismatch; this latter method also draws on Shimer (2005).

Here is an illustrative example of how we calculate the mismatch index within a given skill sector in a given year. We label each 6-digit SOC code as belonging to either the low-, middle-, or high-skill sector, as described above. We count the number of unemployed individuals in each 6-digit SOC code in the CPS in a given time period, and aggregate up to the 2-digit SOC level, but only include the numbers of unemployed workers in the 6-digit occupations that lie in the given skill sector. To take management occupations (SOC 11) as an example, the number unemployed in the "high-skill" sector within management occupations would be the sum of all unemployed individuals across each of the "high-skill" 6-digit SOC codes housed under the management umbrella. We perform similar calculations to obtain the number of vacancies by 2-digit SOC code by skill sector. These values, combined with the model parameters referred to above, are then placed into the formula provided in equation (3) in the main text. A graphical illustration of the partitioning and aggregation process is given in Figure A5.

Separately, in order to calculate the mismatch index within a given 2-digit SOC code , we pull the number of unemployed for each 3-digit SOC code beneath the 2-digit umbrella from the CPS, and also pull the number of vacancies for the same set of 3-digit occupations from the normalized BGT dataset. Staying within a 2-digit SOC group, we then place these values (and the appropriate model parameters) into the mismatch index formula. See Figure A6 for a graphical illustration.

Figure A5. Constructing Skills Mismatch Indexes by Skill Sector

1. Measure the number of unemployed workers by occupation at the six-digit SOC level.

2. Categorize six-digit occupations as low/middle/high skill using 2005-07 ACS.

3. Disaggregate each two-digit SOC into low /middle/high components using the classified six-digit SOCs within it.

Healthcare Practitioner & Technical Occupations

29-0000 (100 U)

29-1064 OBGYN (high-skill) = 30 unemployed

29-1051 Pharmacists (middle-skill) = 50 unemployed

29-9012 OSHA technician (low-skill) = 20 unemployed

4. Re-aggregate number of unemployed into two-digit occs by skill sector:

Healthcare Practitioner & Technical Occs: High-Skill
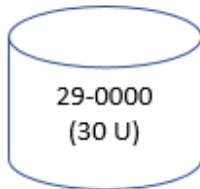
29-0000 (30 U)

Healthcare Practitioner & Technical Occs: Middle-Skill
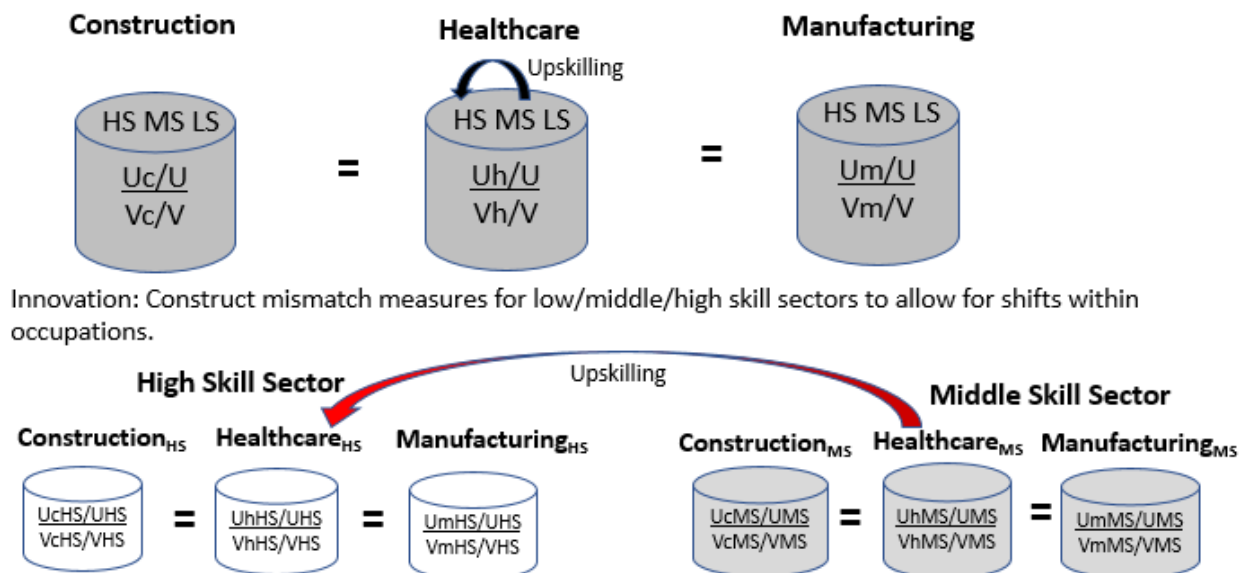
29-0000 (50 U)

Healthcare Practitioner & Technical Occs – Low-Skill

29-0000 (20 U)
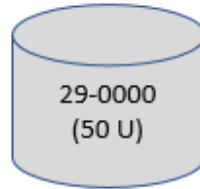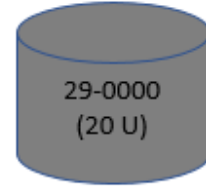
Source: Authors' illustration.

Figure A6. Assessing Changes in Mismatch over the Business Cycle by Skill Sector

Source: Authors' illustration.

III. Robustness Checks

A.   Equilibrium Labor Market Responses to Upskilling by Skill Sector

Although the skill and/or education requirements contained in online job postings are an indicator of employer demand, the advertisements may represent idealized attributes for job candidates, rather than strictly binding qualifications. If employers were even partly successful in upskilling their jobs—in the sense of hiring a higher share of college graduates into a given occupation—we would expect the average education level of new hires to increase relative to that of continuing employees. Since persistent upskilling was more prevalent within the high-skill sector, we would expect any increase in the relative education level of new hires to be most pronounced within that sector.

As a robustness check to test whether rising education requirements were in fact binding, we use the longitudinal component of the CPS to plot the ratio of the average education level of new hires to that of continuing employees by skill sector for the period from 2005 through 2017—including 2008 and 2009 because the data from those years are not missing from the CPS. Educational attainment is coded using discrete numerical values that are increasing in attainment.

Ratios greater than one indicate that the average education level of new hires exceeds that of continuing employees, and values less than one indicate the reverse.[30]

Figure A7 presents the trend in relative education levels across new hires, continuing employees, and the unemployed, based on the CPS data. Does the education level of new hires increase relative to incumbents (e.g., do employers get the type of better-skilled employees they were demanding?). Among high-skill occupations, where much of the persistent upskilling occurred, the relative average education level among new hires experienced a modestly increasing trend over the time period, notwithstanding some fluctuations, with the ratio increasing from roughly 0.91 in 2005 to 0.95 in 2017.[31] Among middle-skill occupations, where much of the temporary upskilling occurred, the average education level of new hires relative to continuing employees increased beginning in 2008-2009 and upskilling remained elevated through 2014, but then these gains reversed by 2017, suggesting that employers were temporarily able to hire more-educated workers during the Great Recession and even well into the recovery. In contrast, in the low-skill sector, where most occupations experienced no upskilling, the education level of new hires relative to continuing employees was flat during this same period. These patterns suggest that the increased education requirements reflected more than mere aspirations on the part of employers, and that firms were at least partly successful in meeting their demand for more highly educated employees.

B.   Wage and Employment Growth

The evidence that we have presented thus far suggests some connection between the persistence in rising employer skill requirements and stable or rising mismatch, potentially contributing to reduced matching efficiency in the aggregate labor market and therefore to the outward shift in the Beveridge curve. However, the sluggish wage growth observed during most of the recovery period seems inconsistent with the mismatch hypothesis (Rothstein 2012; Abraham 2015). To explore this further, we take a cursory look at wage growth by 2-digit SOC codes across occupations based on their upskilling behavior. Again, we use 2010 through 2015

---

[30] Please see the appendix for further details of how these measures are constructed from the Current Population Survey.

[31] Note that as new hires are disproportionately younger and tend to be hired into lower-ranked jobs, it is not surprising that their average education levels are lower than for continuing employees, who may increase their education levels over time while on the job.

as our period of reference, since the US unemployment rate had returned to its pre-recession levels by the end of 2015. Table A2 reveals that six out of the 10 persistent upskilling occupations—and three out of four of the "pure upskilling" occupations—experienced wage growth between 2010 and 2015 that was faster than in the aggregate. Among the non-upskilling group, only two out of nine occupations had above-average wage growth.

Moreover, Table A2 also shows that three out of the five pure upskilling occupations also experienced employment growth that was greater than in the aggregate during the same time period 2010–2015. However for the non-upskilling occupations, employment growth, by and large, was negatively correlated with wage growth. Prior research suggests that if mismatch is present, then employment growth will be positively correlated with wage growth (Abraham 2015). One reason for this could be that mismatch arises when some occupations have very high V/U ratios and others have very low ones. The occupations with high V/U ratios will need to raise wages by a significant amount to get more employees, because their labor markets are very tight, whereas the occupations with low V/U ratios will have low employment growth (low demand) since they can fill their vacancies by only offering modest wage increases. Thus, the employment and wage growth trends reported here are consistent with the presence of mismatch in occupations that exhibited persistent upskilling in the high-skill sector.

Figure A7. Relative Education Levels of New Hires v. Continuing Employees by Skill Sector, 2005-2017
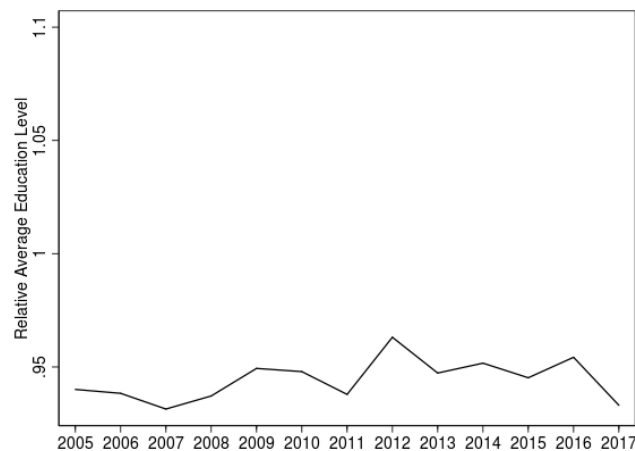
A. High-Skill Occupations
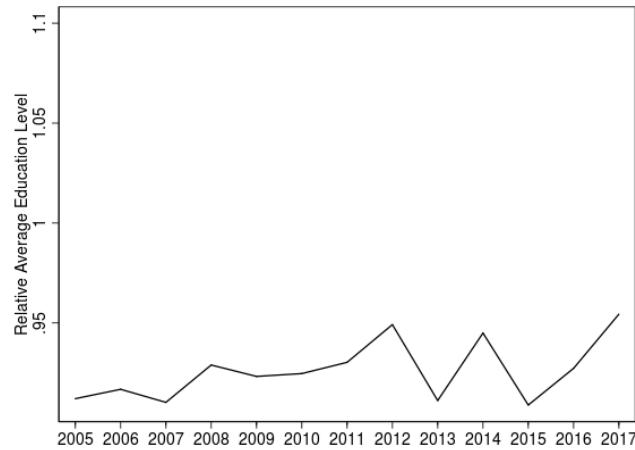
B. Middle-Skill Occupations
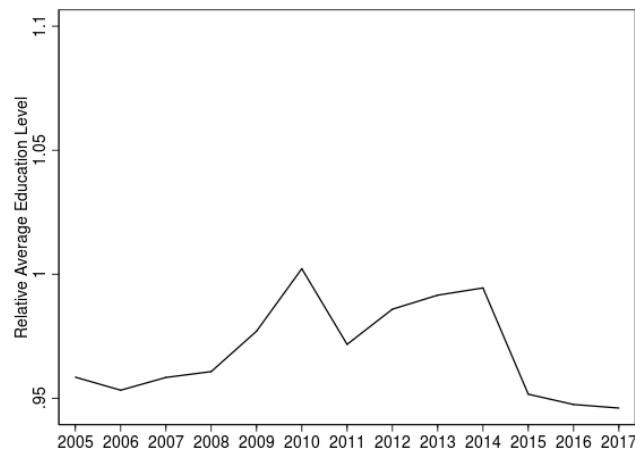
C. Low-Skill Occupations

Source: Authors' calculations using IPUMS-CPS data and IPUMS-ACS data.
Note: Low-skill occupations are defined as those in which at least 40 percent of employees have a high school education or less according to the 2005-07 American Community Survey. High-skill occupations are defined as those in which at least 40 percent of employees have a four-year college degree or greater according to the 2005-2007 American Community Survey. Middle-skill occupations are all other occupations that satisfy neither of these criteria.

67

Table A2. Employment and Wage Trends by Upskilling Trends for 2-Digit Occupations, 2010-2015

Occupation Group	Education Level of Incumbent Workers	Percent of 3D Occupations that Upskilled	Upskilling Type	Employment Growth (Percent)		Real Median Wage Growth (Percent)	
	2005 to 2007	2007 to 2010	2007 to 2010 to 2017	2007 to 2010	2010 to 2015	2007 to 2010	2010 to 2015
All Occupations	-----	-----	-----	-5.40	8.50	3.03	-1.13
Management Occupations	HIGH	100.00	Persistent	0.32	15.18	3.53	-0.36
Business and Financial Operations Occupations	HIGH	100.00	Persistent	1.25	15.46	3.81	0.12
Computer and Mathematical Occupations	HIGH	100.00	Persistent	2.90	21.96	2.04	2.12
Architecture and Engineering Occupations	HIGH	100.00	Persistent	-7.26	7.37	4.25	0.64
Community and Social Service Occupations	HIGH	100.00	Persistent	6.03	3.73	1.08	-1.14
Education, Training, and Library Occupations	HIGH	83.62	Persistent	1.70	1.00	2.63	-4.48
Life, Physical, and Social Science Occupations	HIGH	81.32	Persistent	-15.22	7.67	1.20	-1.84
Healthcare Practitioners and Technical Occupations	HIGH	68.79	Persistent	6.82	9.19	2.75	-1.04
Arts, Design, Entertainment, Sports, and Media Occupations	HIGH	67.08	Persistent	-2.53	7.40	2.22	-0.47
Legal Occupations	HIGH	61.57	Persistent	-0.59	7.02	2.24	-3.12
Sales and Related Occupations	MIDDLE	37.86	Temporary	-6.24	7.62	-1.78	-2.66
Protective Service Occupations	MIDDLE	36.31	Temporary	3.24	5.14	4.64	-4.88
Office and Administrative Support Occupations	MIDDLE	23.87	Temporary	-7.59	1.59	1.53	-0.10
Healthcare Support Occupations	MIDDLE	0.00	None	9.31	0.68	-0.62	0.99
Construction and Extraction Occupations	LOW	8.80	None	-24.38	7.99	2.26	0.02
Installation, Maintenance, and Repair Occupations	LOW	8.17	None	-8.56	9.03	2.25	-1.42
Production Occupations	LOW	5.70	None	-18.83	10.16	3.04	-1.66
Personal Care and Service Occupations	LOW	0.74	None	2.57	25.76	-0.15	-2.15
Food Preparation and Serving Related Occupations	LOW	0.00	None	-2.19	14.05	4.67	-3.56
Building and Grounds Cleaning and Maintenance Occupations	LOW	0.00	None	-5.19	5.54	1.54	-1.91
Farming, Fishing, and Forestry Occupations	LOW	0.00	None	-8.92	11.32	0.97	2.44
Transportation and Material Moving Occupations	LOW	0.00	None	-11.23	11.57	3.26	-2.07

Source: Online job vacancy data provided by Burning Glass Technologies, unemployment and employment and wage data from the Occupational Employment Statistics.