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

Mary A. Burke\textsuperscript{a}
Alicia Sasser Modestino\textsuperscript{b}
Shahriar Sadigh\textsuperscript{c}
Rachel Sederberg\textsuperscript{d}
Bledi Taska\textsuperscript{e}

This version: October 2018

Abstract: In this paper, we test the hypothesis that changes in skill requirements within some occupations have reduced matching efficiency within some classes of jobs and contributed to the outward shift in the Beveridge Curve since 2007. We find evidence that supports this hypothesis from a few different sources. First, we extend the framework developed by Sahin et al. (2014) to measure labor market mismatch, decomposing measures of mismatch by the education requirements of the underlying occupations. We find that between 2007 and 2017, high-skill occupations displayed higher mismatch rates than did either middle-skill or low-skill occupations, and high-skill occupations experienced a smaller decline in mismatch rates during the recent recovery. We also describe movements in employer education requirements over time using a novel database of 87 million online job posting aggregated by Burning Glass Technologies. These data offer evidence of relatively permanent upskilling within high-skill occupations, where such upskilling may reflect changes in the technologies that complement such jobs. Together the results suggest that the persistently lower matching efficiency among high-skill jobs may be an artifact of increasing skill requirements within jobs, rather than the result of changes in the relative demand for occupations. Our results imply that workers who accumulate specific human capital targeting a distinct labor market may be chasing a moving target, even as they achieve a high level of educational attainment. Furthermore, our findings suggest that workforce development strategies and education policies could benefit from the use of real-time labor market information on employer demands in the current environment.

\textit{JEL classifications:} D22, E24, J23, J24, J63.

\textit{Key Words:} labor demand, skills, vacancies, unemployment, firm behavior.

\textsuperscript{a}Federal Reserve Bank of Boston, 600 Atlantic Avenue, Boston MA 02215. \texttt{m.burke@bos.frb.org}
\textsuperscript{b}Northeastern University, School of Public Policy and Urban Affairs, 310 Renaissance Park, 360 Huntington Avenue, Boston MA 02115. \texttt{a.modestino@neu.edu} (Corresponding Author)
\textsuperscript{c}Amazon, Seattle, WA. \texttt{shahrias@amazon.com}
\textsuperscript{d}Northeastern University, Department of Economics, 409 Holmes Hall, 360 Huntington Avenue, Boston MA 02115. \texttt{r.sederberg@husky.neu.edu}
\textsuperscript{e}Burning Glass Technologies, One Lewis Wharf, Boston MA 02110. \texttt{btaska@burning-glass.com}

Funding: This work was supported by a seedling grant from the NULab for Texts, Maps, and Networks at Northeastern University.
I. Introduction

The persistent weakness of the U.S. labor market during the period following the Great Recession remains poorly understood. As of 2012, two years after the official end of the Great Recession, the unemployment rate still hovered around 8 percent, despite an increase in employer-reported vacancies. Even more puzzling has been the low labor force participation rate which has persisted through 2018, despite the U.S. reaching a historically low unemployment rate of 3.9 percent. Prior research indicates that employers increased skill requirements within occupations during the Great Recession, such that unemployed workers may no longer qualify for the jobs that they previously held (Modestino, Shoag, and Balance 2016). Although some of the increase in skill requirements was cyclical, as much as two-thirds of the increase appears to have persisted through the recovery, possibly leading discouraged workers to drop out of the labor market (Hershbein and Kahn 2018, Zago 2015). Taken together, these trends could be evidence of labor market mismatch, a structural decrease in matching efficiency between unemployed workers and job openings due to increasing skill requirements within occupations.

Indeed, the outward shift in the relationship between unemployment and vacancies, known as the Beveridge curve, has highlighted the need to focus not just on the number of vacancies, but on their composition and skill requirements as well (Diamond and Sahin 2014). During the Great Recession, the Beveridge curve shifted out such that vacancy rate was higher for any given level of the unemployment rate (Figure 1). This outward shift in the Beveridge curve persisted even as the labor market recovered. As of July 2018, despite the unemployment rate falling below its pre-recession trough, the vacancy rate remains higher than it was in 2007.

Several explanations for this shift have been proposed, each with potentially different policy implications. Some have interpreted the shift as a deterioration in the matching/hiring
process in the economy, such that idle workers may be seeking employment in sectors other than those with available job vacancies. For example, Sahin et al. (2014) measure mismatch between vacancies and unemployed workers across industries, occupations, and geographies and find that mismatch can potentially account for one-third of the increase in the unemployment rate during the Great Recession. However, the sluggish wage growth observed during the recovery period, even within industries and occupations with relatively strong demand in the United States, would suggest little or no role for labor market mismatch (as shown, for example, by Rothstein 2012 and Abraham 2015). Indeed, several previous papers have argued that weak aggregate demand affecting a broad swath of industries and occupations during the Great Recession offers a more convincing explanation for the outward shift of the Beveridge Curve than does skills mismatch or other structural factors (Ghayad and Dickens 2012, Lazear and Spletzer 2012, Rothwell 2012, Haltiwanger, Davis, and Faberman 2012, Carnevale et al. 2012). Still others have sought to explain the shift in the Beveridge curve by exploring the roles of reduced employer recruiting efforts (Haltiwanger et al. 2012, Davis et al. 2012), increased uncertainty (Barnichon et al. 2012, Daly et al. 2012), pre-recession trends in matching efficiency (Hall and Schulhofer-Wohl 2015), 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 Sahin 2014), among other factors.

Nevertheless, business leaders and policy makers, especially at the state and local level, continue to cite a “skills gap” as a major obstacle to current and future job growth, particularly within certain sectors of the economy such as manufacturing and healthcare. According to a joint study by the Manufacturing Institute and Deloitte from 2015,¹ some 54 percent of surveyed

¹“The Skills Gap in Manufacturing: 2015 and Beyond.”
http://www.themanufacturinginstitute.org/~/media/827DBC76533942679A15EF7067A704CD.ashx
manufacturing executives perceived that there was a high-to-severe shortage of skilled production workers, and based on the survey it was forecasted that close to 60 percent of expected manufacturing vacancies between 2015 and 2025 would go unfilled due to skills gaps. Additional employer surveys and numerous media reports have reinforced the suggestion that a lack of skilled workers has made it difficult to fill some jobs during the economic recovery, leading to slower than expected improvement in the labor market.2

One potential explanation for the disconnect between the anecdotal evidence and academic findings is that the skill demands of employers are changing within occupations, such that even workers with pre-recession experience in a given occupation are not fully-equipped to fill newer jobs in the same field during the recovery. There is at least anecdotal evidence that job descriptions in some industries have been revised in response to recent structural changes to include new tasks requiring more advanced skills. For example, the Affordable Care Act, by reducing the reimbursement rates for a host of medical procedures and services, allegedly raised the skill requirements for nurses and physicians’ assistants, as services formerly rendered by doctors are increasingly pushed down the skill hierarchy.3 As another example, innovative processes such as additive manufacturing require workers to master computer-aided-design (CAD) software and 3D printing in order to produce parts formerly made using only analog technologies.4 Such structural changes within occupations might help explain employers’ claims

---


3 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/

4 See https://www.ge.com/additive/additive-manufacturing.
that they can’t find qualified workers for some jobs, despite economists finding that the distribution of job-seekers across all occupations appears, at least recently, roughly in balance with the occupational distribution of job vacancies.

In this paper, we examine the hypothesis that permanent upskilling within some occupations has contributed to decreased matching efficiency for those occupations, and therefore may help to explain why Beveridge Curve among those occupations continues to exhibit an elevated vacancy rate relative to the unemployment rate, even though the unemployment rate has fallen back to below its pre-recession level. In the first part of the paper, we use data from Burning Glass Technologies covering nearly all online job postings (82.5 million total) in the U.S. between 2007 and 2017, to explore changes in education requirements within occupations to determine which sectors of the labor market (low-, middle-, or high-skill) engaged in permanent versus temporary upskilling over the business cycle. We find that high-skill occupations are more likely to have increased education requirements during the recession and more likely to have maintained those requirements during the recovery, while middle- and low-skill occupations typically exhibited only temporary upskilling. Furthermore, we find some evidence suggesting that the increased education demands of employers were at least partly binding in equilibrium. In particular, in the Current Population Survey we observe that between 2007 and 2017 the average level of educational attainment of new hires increased relative to that of continuing employees for workers in the high-skill occupations, while for workers in low-skill and middle-skill occupations the relative education level of new hires was either flat or exhibited only a temporary increase.

We then use the Burning Glass data to describe changes in specific skill requirements (e.g., baseline, specialized, and software skills) within occupations to better understand which
types of skills were driving the changes in education requirements and whether they might reflect changes in technology or other structural forces. We find that most of the permanent increase in skills within the high-skill sector of the labor market stems from software skills while changes in baseline skills (e.g., project management) and specialized skills (e.g., information security) were more temporary in nature. We confirm that wages were rising faster for jobs requiring these software skills within high-skill occupations, suggesting that employers value workers with these skills rather than simply using them as a screening device.

In the second part of the paper, we extend the mismatch index developed by Sahin et al. (2014) to explore labor market mismatch for different skill sectors of the labor market. Specifically, we construct mismatch indices by education level (low/medium/high) to determine whether there is heterogeneity in mismatch across labor markets segmented by their education or “skill” type. Importantly, we also contrast changes in matching efficiency, as measured by the Beveridge Curve, between occupations that experienced temporary upskilling during and after the Great Recession from those that experienced permanent upskilling. We find that mismatch varies across different skill sectors of the economy, in terms of both its average rate (and its contribution to unemployment) in recent years as well as its persistence during the recovery period. Grouping occupations by education level of employed workers prior to the Great Recession, we find that the mismatch index is greater for high- versus middle- and low-skilled sectors during the period 2007 through 2017. Moreover, while the mismatch index increased within all sectors during the Great Recession, it subsequently receded for low- and middle-skill jobs yet remained elevated for jobs in the high-skill sector of the labor market.

Examining separate Beveridge curves by skill sector, the movements in the respective curves are consistent with the dynamics of skill requirements observed within occupations. That
is, the Beveridge curve for low- and middle-skill jobs shifted out during the recession but have since largely reverted to their pre-recession positions, while the Beveridge curve for high-skill occupations exhibited a more persistent outward shift since 2007. This outward shift in the Beveridge curve for high-skill occupations suggests that there has been a persistent decline in matching efficiency within that segment of the labor market, a decline which could reflect the effects of upskilling within high-skill occupations and/or increased mismatch across high-skill occupations. However, the reasons for the apparent decline in matching efficiency may be various, and will be discussed below.

The finding that mismatch varies across different skill sectors of the economy—not only in levels but also in terms of whether increases in mismatch during the Great Recession were temporary or long-lasting—can shed light on the dynamics of the labor market, particularly over the course of the business cycle. For example, aggregate measures of mismatch may understate the difficulty of filling vacancies in certain skill sectors—particularly, at present, occupations for which a bachelor’s degree or higher is required. Our finding that the demand for skills within low- and middle-skill occupations is more sensitive to the business cycle also explains why one observes higher-skilled workers moving down the “career ladder” during recessions (Zago 2015, Beaudry et al. 2013, Carnevale et al 2012). Finally, the observation that high-skilled occupations were subject to relatively persistent (i.e. structural) increases in mismatch and persistent upskilling may help explain findings from other studies showing that, since roughly 2002, the college wage premium has stalled for those with only a bachelor’s degree but has grown for workers with advanced degrees (Acemoglu and Autor 2011 and Acemoglu and Autor 2012).

Our findings also speak to debates about workforce development strategies and related educational policies, an arena which could benefit from the use of real-time labor market
information on the demand for and the supply of specific skills. For example, prior research shows that more recent business cycles have been characterized by shifts in employment across skill sectors that have persisted during recovery periods (Jaimovich and Siu 2014). Similarly, desired qualifications for low- and middle-skill jobs may now increase temporarily over the business cycle, suggesting that workers displaced from these skill sectors during a recession may experience longer spells of unemployment than in earlier decades. If so, then a targeted extension of unemployment insurance UI benefits beyond the usual 26 weeks to low- and middle-skill workers may be warranted. As another example, an ability to identify permanent shifts in skill demands within certain occupations could help identify human capital investments with higher long-term payoffs.

The paper proceeds as follows. Section 2 describes the relevant literature on labor market mismatch. Section 3 describes the data and methods that we use to measure changes in the supply and demand sides of the labor market. Using the BGT data, Section 4 examines the demand-side changes in education and skill requirements within occupations over the business cycle for different sectors of the labor market (low-, middle-, or high-skill). Section 5 draws on the Current Population Survey to determine whether rising skill requirements were indeed binding by looking at the education levels of new hires versus continuing workers each skill sector. Section 6 combines both the supply- and demand-side data to generate estimates of labor market mismatch by skill sector over the most recent business cycle and relate these dynamics to movements in the Beveridge curve by skill sector. Section 6 concludes with a discussion of potential policy implications.
2. Related Literature

As noted above, many recent papers have downplayed or dismissed the contribution of structural labor market mismatches to the slow recovery of unemployment from the recession and to the decline in aggregate matching efficiency evidenced by the Beveridge Curve’s outward shift since 2007. The results of Sahin et al. 2014 represent something of an exception in that regard, in that they show that an increase in mismatch, even if temporary, can exert relatively persistent upward pressure on unemployment. However, relatively little work has been done to study the relevance of skills mismatch within occupations for the recent labor market experience in the United States, and formal studies of changes in skill requirements by occupation are few in number. Vaisey (2006) compares the education requirements of jobs (observed using the US Department of Labor’s O*NET database) to the educational attainment of workers employed in the same jobs and finds that the average worker was overqualified for his/her job as of 2002. Also using the O*NET database, Liu and Grusky (2013) find some evidence that some skill requirements—including computer skills and analytic and quantitative skills—increased within jobs since 1979, but the increases are small to modest and, perhaps surprisingly, did not include engineering skills.

Skill-biased technological change has been cited as a factor leading to increased demand for highly-educated workers (performing in highly-skilled occupations) relative to less-educated workers (see, for example, Tinbergen 1974, Goldin and Katz 2008, Acemoglu and Autor 2011). However, the evidence is mixed as to whether the adoption of new technologies raises skill requirements within jobs (see for example Acemoglu 2002, Zicklin 1987, and Keefe 1990).

Cappelli (2014 and 2015) questions whether requirements listed in vacancies are truly binding, and emphasizes the problem of overqualification within occupations rather than
underqualification. He suggests that employers are reducing investments in training and trying to shift the burden of workforce training onto the public sector. In contrast, Bessen (2014) finds wage-based evidence of a gap in computer skills affecting many occupations. He uses growth in 90th percentile wages relative to median wages as an indicator of the value of scarce skills, and finds that occupations involving relatively high computer use, including scientific occupations and healthcare jobs, saw large increases in his relative wage growth measure, while occupations involving less computer use saw smaller relative wage increases.

Zago (2015) studies the relationship between job polarization and skill (i.e. educational) mismatch in the wake of the Great Recession (through 2013). Specifically, he shows that states that experienced greater job polarization during the recession—a loss of routine (middle-skill) jobs in favor of both manual (low-skill) and abstract (high-skill) jobs—also experienced greater educational mismatch, in the sense that workers increasingly walked down the occupational skill ladder relative to their education credentials. The results suggest that, over the longer-run, structural (rather than cyclically induced) job polarization may contribute to upskilling within occupation classes as well as to increased incidence of educational overqualification in some classes of jobs.

Modestino, Shoag, and Ballance (2016) showed that many employers increased their education requirements following the Great Recession. While much of the upskilling they observed was found to be temporary or opportunistic, up to two-thirds of upskilling was found to be persistent, or apparently structural. Building on those findings, the current paper describes upskilling patterns over a longer time period and decomposes them by skill sector, according to whether the job vacancies typically require either low, middling, or high education levels. Going further, the current study identifies large and permanent increases (since 2007) in the demand for
software skills among the set of firms that engaged in permanent education-based upskilling since 2007, whereas firms imposing only temporary increases in education requirements displayed much smaller increases in demand for software skills that were frontloaded between 2007 and 2010 and flat thereafter. Decomposing the Beveridge Curve and mismatch indices by skill level we find evidence that further suggests that high-skilled occupations experienced larger declines in matching efficiency since 2007 than other skill sectors, consistent with the observation that high-skilled occupations were more subject to permanent increases in education and specific skill requirements than other sectors.

3. Data and Methods

Our primary objective is to explore the degree to which upskilling within occupations over the business cycle contributed to labor market mismatch, possibly resulting in an outward shift of the Beveridge curve. To achieve this goal, we will need to examine the dynamics on both the demand and supply side of the labor market as well as how they match up over time. Specifically, we will seek to answer the following research questions:

- How have employer education requirements increased within occupations and in which skill segments (low-, middle-, and high-skill) of the labor market? Are these increases permanent or temporary?
- Have actual skill requirements also increased within occupations? Which sets of skills are now in greater demand?
- Are these skill requirements binding? Do new hires have greater levels of education than continuing workers within the same occupation?
• Do permanent increases in skill requirements mean that unemployed workers are no longer qualified for the jobs they once held? Do the unemployed have lower levels of education than continuing workers within the same occupation?

• Does upskilling within occupations increase labor market mismatch? Does this vary across low-, middle-, and high-skill sectors of the labor market? If so, can this help explain the outward shift in the Beveridge curve?

To test our hypotheses, we will use a variety of methods drawing on multiple data sources to provide a collage of evidence from which we can draw our conclusions. Our analysis will consist of three primary parts. First, we will use the near-universe of online job vacancy data provided by Burning Glass Technologies to examine the demand-side changes in education and skill requirements within occupations over the business cycle for different sectors of the labor market (low-, middle-, or high-skill). Second, we use the Current Population Survey to determine whether rising skill requirements were indeed binding by looking at the education levels of new hires versus continuing workers versus the unemployed within each skill sector. Finally, we combine both the supply- and demand-side data to generate estimates of labor market mismatch by skill over the most recent business cycle and relate these dynamics to movements in the Beveridge curve by skill sector. We describe each of these analyses below in terms of the data and methods used as well as how we divide the labor market into the three skill sectors used throughout the paper. But first, we need to describe the data sources used to measure both the demand and supply sides of the market.

3.1 Data Sources

3.1.1 Online job posting data from Burning Glass Technologies (BGT)
The vacancy data used in this paper is collected by Burning Glass Technologies (BGT), one of the leading vendors of online job posting data. BGT collects detailed information on the more than seven million current online job openings daily from over 40,000 sources including job boards, newspapers, government agencies, and employer sites. 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 text parse each job ad into usable data elements. BGT mines over seventy job characteristics from free-text job postings including employer name, location, job title, occupation, years of experience, level of education required.

Unlike other online job vacancy sources, BGT also parses out other dimensions of skill from the text of the job ad, allowing us to create measures of different types of skill rather than simply relying on education as a proxy for skill. They aggregate this data in several ways. First, they collect each skill from a job posting and then parse it into a canonicalized version to enable improved search and categorization. For example, Python 3.3 and Python 2.7 are both standardized to Python. Because this results in over 16,000 canonicalized skills, BGT then further categorizes each canonicalized skill into skill clusters. For example, algebra and calculus would both be categorized into the math skill cluster. These skill clusters are then further classified into skill families (e.g. math and science are both STEM skills). Finally, skill families are then classified as baseline skills (e.g. leadership), specialized skills (e.g., accounting), and software skills (e.g., Oracle).

The collection process employed by BGT provides a robust representation of hiring, including job activity posted by employers. The process follows a fixed schedule, “spidering” a
pre-determined basket of websites that is carefully monitored and updated to include the most
current and complete set of online postings. BGT has developed algorithms to eliminate
duplicate ads for the same job posted on both an employer website as well as a large job board
by identifying a series of identically parsed variables across job ads such as location, employer,
and job title. In addition, to avoid large fluctuations over time, BGT places more weight on large
job boards than individual employer sites which are updated less frequently.\footnote{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.}

In the database provided by BGT, a snapshot of vacancies is reported monthly and are
pooled over the year without duplication. This data is unique in allowing geographical analysis
of occupation-level labor demand for a variety of skills including education and experience over
time. Using the entire universe of job vacancies collected by BGT, allows us to expand the labor
mismatch analysis to occupational level. The data are available for detailed occupation by
Standard Occupation Code (SOC) down to the six-digit level for 2010 through end of 2015.\footnote{We have aggregated the data to the two-digit occupational categories using the appropriate mappings from the 2010 SOC codes.}

It should be noted that although Burning Glass Technologies consistently applies the
same filtering and de-duplication algorithm across years, even retroactively as improvement are
made, the number of sources scraped may have evolved over time. Figure 2 plots JOLTS
vacancies and BGT ads at the national level. The total count of active vacancies in BGT is
below that in JOLTS although the correlation between the two series is quite strong at 0.82. To
the extent that the trend in online vacancies is similar to that of JOLTS across sectors, our
calculations should not be affected. Yet, there are also specific occupations which are
underrepresented in all on-line job posting data, such as construction jobs, that are not typically
posted online. Because the actual number of vacancies is important for measuring mismatch, we use a normalized data set provided by BGT that re-weights the raw BGT to adjust the total number of postings at the industry level to match the number of monthly vacancies from JOLTS. This reweighting process produces a BGT vacancy data series that is consistent and comparable with the JOLTS and other vacancy data series that use a similar weighting methodology such as HWOL.9

3.1.2 Current Population Survey

On the labor supply side, we make extensive use of the Bureau of Labor Statistics’ Current Population Survey microdata (IPUMS-CPS, Flood et al. 2018) from 2007 through 2017. In particular, to construct our mismatch indexes we use the CPS data to measure the number of unemployed workers by occupation at the six-digit SOC level, and when estimating mismatch-related unemployment 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 and continuing employees, in order to compare relative education levels between those two groups over time by occupational skill sector

3.2 Methods

3.2.1 Defining low-, middle-, and high-skill sectors

We define low-, middle-, and high-skill sectors in two different ways, although the two methods yield very similar results. On the supply side, we use the 2005-07 American Community Survey (ACS) to generate the education distribution for each six-digit occupation

9 See the appendix for details.
(based on SOCs) according to the credentials held by the incumbent workers. We then calculate
the share of low-skill workers (high school graduates or less), middle-skill workers (some
college and associate degree holders), and high-skill workers (bachelor’s degree or higher) by
occupation. High-skill occupations are defined as those where at least 40 percent of the
vacancies require a bachelor’s degree or higher, low-skill occupations are those where at least 40
percent of the vacancies require a high school degree or less, and middle-skill occupations are
those where the education requirements lie somewhere in between or at least 40 percent require
an associate degree or some college. We keep these definitions fixed throughout the study
period.

We use these classifications to estimate mismatch indexes separately by occupational
skill group, at the level of the 2-digit SOC. In particular, using the CPS we obtain counts of the
number unemployed for each six-digit occupation, and then aggregate these up to the two-digit
SOC level, separately by skill sector. That is, rather than labelling a given 2-digit occupation as
either low-, middle-, or high-skill in its entirety, each 2-digit SOC is partitioned into low-,
middle-, and high-skill components, where each component sums just the data (e.g. numbers of
unemployed) from the underlying 6-digit occupations under the given 2-digit umbrella and with
the given skill level. For example, unemployed workers within the two-digit Management
occupation grouping (SOC 2010 = 11) would be allocated across high-, middle-, and low-skill
components according to the educational classification of the six-digit occupations within it.
These numbers of unemployed by 2-digit occupation by skill group enter into the calculation of
the mismatch indexes by skill group, together with vacancies by 2-digit occupation by skill
group that are constructed similarly (as described below).
We perform an analogous exercise on the vacancies or demand-side using the education requirements of vacancies contained in the BGT data. That is, we label each 6-digit occupation as low-, middle-, or high-skill according to the share of vacancies in the occupation that require a given level of education. As before, we generate vacancies by 2-digit occupation by skill level by adding up vacancies among 6-digit occupations within the given skill group and the given 2-digit umbrella, and use these in the calculations of the mismatch index by skill group. This approach represents an innovation over previous research (e.g. Sahin et al. 2014) that applied the same ACS distribution to both the supply- and demand-sides of the mismatch analysis due to a lack of demand-side educational requirements. Using the same distribution for both sides of the market implicitly assumes that the educational requirements of newly created vacancies in a given occupation will mirror the educational levels of incumbent workers, and thereby fails to allow for any upskilling over time. For example, although 60 percent of currently employed administrative assistants do not have a bachelor’s degree, over 60 percent of job postings for administrative assistance require one.

However, using the BGT data to estimate the educational distribution of job postings is not without measurement error. By doing so, we assume that online job postings represent all job vacancies. In addition, a non-trivial share of online job postings list no education requirement, and these cases can occur either for very low-skilled jobs (e.g., dishwasher) or for very high-skilled jobs (e.g., CEO), where in either case the educational credential is not listed because it is understood. As it turns out, the distribution of education requirements among vacancies in the BGT data at the two-digit SOC level is very similar to the distribution of education qualifications of incumbent workers estimated using the ACS during this period.
3.2.2 Measuring demand-side upskilling within occupations by skill sector

We use the BGT data to examine demand-side changes in education and skill requirements within occupations over the business cycle for different sectors of the labor market (low-, middle-, or high-skill). Here, we define the initial low-middle-high skill sectors using the ACS and hold them fixed over the period 2007-17. We then allow education and skill requirements to vary within those skills sectors based on the underlying jobs postings from the BGT data over the business cycle. For example, we compare the change in share of postings within high-skill occupations during the recession (2007-10) and recovery (2010-17) periods.

We then define “permanent” upskillers as those occupations for which there was a 5-percentage point increase in the share of postings requiring a bachelor’s degree between 2007 and 2010 (the average increased observed by Modestino et al. 2016 during this period) and less than a 2.5 percentage point decline in that same share between 2010 and 2017. Temporary upskilling is characterized by a similar run-up in education requirements between 2007 and 2010, but a 2.5 percentage point or greater decline in the share of postings requiring a bachelor’s degree between 2010 and 2017. The “no upskilling” label applies to occupations that experienced less than a 5-percentage point increase in bachelor’s degree requirements between 2007 and 2010. We then compare the extent of permanent, temporary, and no upskilling across our low-, middle-, and high-skill sectors.

To test for actual changes in skill requirements, we make use of the skills data collected by BGT from the text of the job postings. Specifically, we perform a difference-in-difference analysis of changes in the share of postings requiring baseline, specialized, and software skills for permanent versus temporary upskilling occupations over time—both during the recession
versus the recovery. We then delve further into the skills data to examine the most requested skill clusters within the baseline, specialized, and software skill categories.

3.3 Measuring labor market equilibrium responses to upskilling by skill sector

The skill demands and/or education requirements contained in job postings might represent employers’ aspirations for job candidates rather than strictly binding qualifications. If increased education requirements are in fact binding for a given occupation, we would expect to observe an increase in the past decade or so in the average education level of new hires in that occupation, relative to the average education level of continuing employees in the same occupation. We also expect the average education level of unemployed workers in a given occupation to fall behind that of continuing workers when employers engage in upskilling. To test these hypothesis, we identify new hires and continuing employees by 4-digit occupation in the CPS (using a crosswalk from 6-digit SOC codes to the CPS coding system) by exploiting the survey’s panel dimension. The method of matching individuals in the CPS longitudinally are adapted from Madrian and Lefgren (2000) and Hershbein (2017). An individual is deemed a new hire in a given month if between the previous month and the current month the individual made any of three types of employment-status transitions: (1) from not in the labor force to employed, (2) from unemployed to employed, or (3) from employed to employed but with a change in employer. An individual counts as a continuing employee in a given month if they were employed in both the current month and the previous month and experienced no change in employer. (Therefore all new and continuing hires are drawn from the set of individuals observed in both the previous month and the current month.) For each new hire or continuing employee, we observe both the occupation

---

10 One might also include as new hires those who stay employed and with the same employer, but who experience a change of occupation between time periods. However, these types of transitions may give false positives for new hires, for example in the case of being promoted up a career ladder but with no significant change in job description.
and the education level of the individual. Separately for new hires and continuing employees and for each year between 2005 and 2017, we sum the education information up to the level of the skill sector using the relevant observations from all CPS occupations within that skill sector. By assigning a discrete value to each education level—1 for high school or less, 2 for some college or associate’s degree, 3 for bachelor’s degree and 4 for advanced degree—we calculate the average education level among new hires (and, separately, among continuing employees) by skill sector, and then use these values to calculate the relative average education level of new hires compared with continuing employees. In a similar fashion we calculate the average education level of unemployed individuals by skill sector, using the occupation listed in the CPS for the unemployed person’s most recent job.

3.4 Measuring labor market mismatch by skill sector

To measure changes in labor market mismatch across sectors, we extend the methodology developed by Sahin et al. (2014) to construct an occupational mismatch index further into the recovery within each of our three skill sectors (low-, middle-, and high-skill). In this section, we briefly describe the measure of labor market mismatch that we employ, which directly follows the index developed by Sahin et al. (2014). In this framework the labor market is frictional in that each occupation constitutes a distinct labor market. This means, for example, that a given vacancy or a given worker is associated with a single occupation type, and a vacancy in a given occupation can only be filled with a worker associated with the same occupation—there is no movement of workers across occupations over time. The framework implies that the costs to workers of switching occupations and/or the costs to firms of hiring a worker from a different occupation are prohibitively high. This is likely to be a reasonable assumption for
occupations defined at the 2- or 3-digit SOC level for which there is less movement than say, across 6-digit SOC levels.

Following Sahin et al., we construct a mismatch index that quantifies the fraction of hires lost because of misallocation between unemployed workers and vacancies across occupations as derived from a planner’s optimal rule.\(^\text{11}\) Letting \(h_t\) denote the observed aggregate hires and \(h^*_t\) the planner’s hires, simply put the mismatch index captures the fraction \((1 - h_t / h^*_t)\). The number of hires in occupation \(i\) at time \(t\), denoted \(h_{it}\), is governed by a matching process that can be represented as follows:

\[
h_{it} = \Phi_t \varphi_{it} m(u_{it}, v_{it}) = \Phi_t \varphi_{it} u_{it}^{1-\delta} v_{it}^\delta.
\]

According to the above, \(h_{it}\) depends on the level of fundamental frictions in the occupation, \(\Phi_t \varphi_{it}\), together with an underlying matching function, \(m(.,.\))\). The frictions, also referred to as matching efficiency, contain a common aggregate component, \(\Phi_t\), and an occupation-specific component, \(\varphi_{it}\). The matching function, based on empirical evidence is assumed to follow a Cobb-Douglas form, such that hires are increasing (and concave) in both the number of unemployed workers searching in that occupation at that time, \(u_{it}\), and the number of vacancies, \(v_{it}\), posted for that occupation at that time. The Cobb-Douglas parameter \(\delta \in [0,1]\) represents the vacancy elasticity of hires. We assume that the vacancy elasticity is common across occupational labor markets. Summing across these markets, the aggregate number of hires in the economy at time \(t\) can be written as:

\(^{11}\) Mismatch can also be measured across industries. Throughout our discussion we use the example of occupations, but “industries” could be substituted for occupations throughout the analysis with no loss of generality.
\[ h_t = \Phi_t u_t^{1-\delta} v_t^{\delta} \left[ \sum_{i=1}^{I} \varphi_{it} \left( \frac{u_{it}}{u_t} \right)^{1-\delta} \left( \frac{v_{it}}{v_t} \right)^{\delta} \right] \]  

(2)

In this framework, the total level of hires at each date is optimized subject to the matching frictions imposed by each market. The optimal planner’s solution moves the unemployed across sectors to allocate more unemployed workers to markets with higher matching efficiencies and higher vacancies. One can derive the optimal number of hires that arises under the planner’s solution (see Sahin et al. 2014 for details) and express the basic mismatch index, now denoted \( M_t \), as follows:

\[ M_t = 1 - \frac{h_t}{h^*_t} = 1 - \sum_{i=1}^{I} \left( \frac{\varphi_{it}}{\varphi_t} \right) \left( \frac{v_{it}}{v_t} \right)^{\delta} \left( \frac{u_{it}}{u_t} \right)^{1-\delta} \]  

(3)

In the above, \( \varphi_t \) is a CES aggregator of the occupation-level matching efficiencies weighted by their respective vacancy shares. This baseline measure of the mismatch index is defined as the fraction of hires that are lost because of an inefficient allocation of unemployed workers across labor markets by occupation.\(^{12}\) By better allocating the same number of unemployed, the planner could increase the aggregate number of hires. The shortfall between the actual number of hires (in the presence of mismatch) and the optimal number of hires constitutes the direct effect of mismatch. Mismatch further reduces hires due to a feedback effect, but we ignore this latter effect in our calculations.\(^{13}\).

\(^{12}\) Dickens (2010) and Lazear and Spletzer (2012) use an alternative index proposed by Mincer (1966) that only quantifies the number of job seekers searching in the wrong sectors.

\(^{13}\) The feedback effect that further reduces aggregate hires occurs because, in the presence of mismatch, the unemployment rate is elevated relative to the optimum unemployment rate, and the higher unemployment rate further depresses the job-finding rate per unemployed worker.
Sahin at al. (2014) demonstrate three useful properties of the index. First, $M_t$ by construction varies between zero (no mismatch) and one (maximal mismatch). Second, the index doesn’t respond to aggregate shocks to the total number of vacancies or the total number unemployed that don’t alter the distribution of vacancies or unemployed across occupations. Third, $M_t$ is increasing in the level of disaggregation, suggesting that the degree of mismatch observed at the three-digit SOC level should be larger than that observed at the two-digit level. To compute the mismatch indices, we estimate the parameters of the model described above for the vacancy elasticity, the market-specific matching efficiencies $\varphi_{lt}$ and the aggregate level matching efficiency $\Phi_t$. We follow Sahin et al. (2014) and include the details of these estimations in the appendix.

With some further assumptions we can use the mismatch index to develop an estimate of the amount of excess unemployment that can be attributed to mismatch. Still following Sahin et al. (2014), mismatch unemployment can be thought of as the difference between actual unemployment and the level of unemployment that would arise under zero mismatch. The concept is similar to that found in the literature on misallocation and productivity (Lagos 2006; Restuccia and Rogerson 2008; Hsieh and Klenow 2009; Moll 2011; Jones 2013). To calculate the counterfactual unemployment level in the absence of the occupational labor mismatch, we follow Shimer (2005). See the appendix for details of the estimation.

It is important to point out a few caveats and limitations associated with the approach described above. First, this methodology still yields a measure of mismatch across sectors—defined by the jointly observable characteristics of job vacancies and unemployed job seekers—not within sectors. As a result, concluding that mismatch plays a small role at the level of two-digit occupations does not necessarily rule out its importance at the three- or six-digit level.
However, we do calculate separate mismatch indices for each skill sector to gain some insight into the degree of labor market mismatch within groups of occupations.

Second, the measure of mismatch used in this paper captures the sectoral misallocation between job vacancies and unemployed job seekers, excluding employed workers who search on the job and excluding non-labor force participants who might want a job even if they are not actively seeking one. However, Sahin et al. (2014) verify that mismatch between vacancies and unemployment behaves very similarly to an index that also includes, among the job seekers, employed workers who report to search on the job. On the other hand, Hall and Schulhofer-Wohl (2015) find that estimates of matching efficiency depend on whether job-to-job seekers and NILF job-seekers are included.

Finally, the mismatch indices we construct enable us to determine the implications for the hiring process of the allocation of workers and vacancies to different occupations, separately by broad occupation skill grouping, but they do not capture more nuanced sources of misallocation, such as those related to the costs of retraining or migration, relative wage rigidity, risk-aversion, imperfect insurance, or government regulations. Given that the benchmark planner’s allocation assumes that there is costless mobility between sectors, our calculations of the impact of mismatch on hiring and unemployment are likely to represent upper bounds.

4. Demand-Side Upskilling within Occupations by Skill Sector

We know from previous research that employers increased skill requirements within occupations during the Great Recession—most notably requiring a bachelor’s degree or higher. About one-third of this upskilling was cyclical in response to the increased supply of workers and was reversed as the labor market recovered (Modestino et al 2016). Yet as much as two-thirds of the upskilling during the recession was structural (Hershbein and Kahn 2018). Given that the skill demands of employers are changing within occupations, it may be the case that
even workers with experience in a given occupation are not fully-equipped to fill newer jobs in
the same field, possibly resulting in a mismatch between supply and demand within occupations.

In this section, we exploit the detailed information in the BGT data to understand the
observed heterogeneity in upskilling over the business cycle by education group. Specifically,
which skill segment of the labor market (low/middle/high) experienced permanent (structural)
versus temporary (cyclical) upskilling over the last business cycle? To explore this question, we
first categorize each 3-digit occupation as having experienced either “permanent upskilling,”
“temporary upskilling”, or “no upskilling” between 2007 and 2017, and then measure the share
of occupations within each low/middle/high-skill category that fall into these categories. Recall
that permanent upskilling is characterized by a 5-percentage point increase in the share of
postings requiring a bachelor’s degree between 2007 and 2010 and less than a 2.5 percentage
point decline in that same share between 2010 and 2017. Temporary upskilling is characterized
by a similar run-up in education requirements between 2007 and 2010, but a 2.5 percentage point
or greater decline in the share of postings requiring a bachelor’s degree between 2010 and 2017.
The “no upskilling” label applies to occupations that experienced less than a 5-percentage point
increase in bachelor’s degree requirements between 2007 and 2010.

Table 1 shows that across all occupations, just shy of 40 percent qualify as permanent
upskillers, close to 19 percent were temporary upskillers, and almost 42 percent engaged in no
upskilling. However, permanent upskilling was far more prevalent within the high-skill
occupation group—at 80 percent—than in the other groups, whereas temporary upskilling was
most common among middle-skill occupations and lack of upskilling was most likely to occur
among low-skill occupations. Splitting the data differently, Table 2 shows that high-skill
occupations were sharply overrepresented among the set of all occupations classified as
permanent upskillers, whereas low-skill occupations were overrepresented among the non-upskillers. Figure 3 shows the trend in education requirements across low-, middle-, and high-skill occupations in the aggregate and confirms that permanent upskilling had a greater impact in the high-skill sector of the labor market. During the recession, the share of postings requiring a bachelor’s degree increased across all three skill groups, although much more so for middle- and high-skill occupations. Yet during the recovery, this share decreased for low- and middle-skill occupations while remaining elevated and even increasing among high-skill occupations.

While these trends are suggestive, increases in education requirements among permanent upskillers may not reflect structural changes in the actual skills required by the job, but instead may have occurred in response to a permanent increase in the supply of workers with bachelor’s 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 3 performs a difference-in-difference analysis of changes in the education and skill requirements for permanent versus temporary upskilling occupations over time—both during the recession versus the recovery. Between 2007 and 2010 the share of postings requiring a bachelor’s degree increased significantly within both the permanent- and temporary-upskilling groups (although by a larger margin in the former group). However, between 2010 and 2017 the permanent upskilling group exhibited a further increase in the share requiring a bachelor’s degree—by nearly 4 percentage points—compared to a drop of 12.6 percentage points in the temporary-upskilling group.

What’s even more striking is that a similar pattern holds in an analogous difference-in-difference analysis using BGT categories of measurable skill requirements from the individual job postings, rather than educational credentials. These categories include baseline skills (e.g.,
communication), specialized skills (e.g., accounting), and software skills (e.g., Oracle). Similar to the bachelor’s degree requirement, all three types of skills became more prevalent among postings during the recession period for both permanent- and temporary-upskilling occupations. Between 2007 and 2010 both permanent upskillers and temporary upskillers became significantly more likely to require elements from each of the baseline skills group, the specialized skills group, and the software skills group. Although there were no reversals of these types of skill requirements during the recovery period, only permanent upskilling occupations experienced a statistically significant increase in the share of postings requiring baseline, specialized, and software skills between 2010 and 2018. However, even among temporary upskillers the share of postings requiring either baseline skills or specialized skills appears higher in 2017 than in 2010 based on point estimates. In contrast, the share of postings requiring software skills was flat among temporary upskillers but continued to increase among permanent upskillers. As a result, software skills were the only type of skill for which there is a significant difference between permanent and temporary upskilling occupations, suggesting that permanent increases in education requirements may be linked to changes in technology. Figure 4 shows the trend in skill requirements across low-, middle-, and high-skill occupations in the aggregate and confirms that while all three groups saw similar increasing trends for baseline and specialized skills, the share of postings requiring software skills was relatively flat for low-skill occupations but rising for mid- and high-skill occupations.

To further explore the nature of the skills required by postings for different skill sectors of the labor market, we delved more deeply into the BGT data and examined the change during the recovery in specific skill clusters within the baseline, specialized, and software categories. Figure 5 shows the initial share of postings requiring a particular baseline skill cluster in 2010
versus the change in the share of postings requiring that skill cluster between 2010-17. While there is some overlap across low-, middle-, and high-skill occupations in baseline skill clusters such as communication and computer literacy, the increase in the share of postings requesting these skills is highest for the high-skilled occupations and lowest for the low-skill occupations. Perhaps more interestingly are the baseline skill clusters that differ across skill sectors. For example, high-skill occupations are more likely to require research, planning, writing, and problem solving compared to middle-skill occupations which are more likely to require organizational skills and being detail oriented versus low-skill occupations which are more likely to require physical abilities.

A similar pattern emerges for specialized skill clusters, as shown in Figure 6. High-skill occupations tend to require a different mix of skills and at a greater initial frequency with a more rapid increase during the recovery. For example, high-skill occupations are more likely to require teaching, budget management, and business strategy compared to middle-skill occupations that are more likely to require scheduling and retail industry knowledge versus low-skill occupations that are more likely to require food and beverage service, equipment repair and maintenance, and material handling. We observe this differentiation even within a particular industry such as healthcare. Among high-skill occupations, behavioral health and ER and intensive care are most frequently required compared to middle-skill occupations that are more likely to require basic patient care and low-skill occupations that are more likely to require basic living activities support.

Greater differentiation across low-, middle-, and high-skill sectors is observed among software skill clusters. In general, software skills are much less likely to be required for low- and middle-skill jobs with no particular skill cluster being requested for even 1 percent of
postings. The software skills that are required for low- and middle-skill jobs are often very
generic such as knowing Microsoft windows or some type of productivity software. In contrast,
high-skill jobs are more likely to require software skill clusters, at a higher frequency, and with
greater variety including SQL programming, statistical software, C and C++, Java, architectural
design, and integrated development environments. As seen in Figure 7, here again we find
suggestive evidence that technology may be driving permanent upskilling among high-skill
occupations, particularly those that use specialized software packages (e.g., architectural design)
or for which new software can diffuse rapidly, changing the nature of the worker’s tasks (e.g.
EPIC medical record technology software).

5. Equilibrium Labor Market Responses to Upskilling by Skill Sector

Just because employers demand more education or skill, it does not necessarily mean that
demand are binding, particularly in the short-run if there are constraints on the supply
of workers with certain credentials or experience. We test whether increased education
requirements were in fact binding in equilibrium by comparing the average education levels of
new hires relative to that of continuing employees, separately by occupation skill sector, between
2005 through 2017. 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
education levels of new hires to increase relative to that of continuing employees. Since we
observed that upskilling was more prevalent and persistent within the high-skill sector, we would
expect any increase in the relative education level of new hires to be most pronounced for that
set of occupations. As described above, we use the CPS microdata to test this prediction. Figure
8 shows that among high-skill occupations, where much of the permanent upskilling occurred,
this was indeed the case. Among middle-skill occupations where much of the temporary
upskilling occurred, the educational levels of new hires relative to continuing employees rose and then fell, suggesting that employers were able to temporarily hire better candidates during the Great Recession. In contrast, among low-skill occupations where most occupations experience no upskilling, the education levels of new hires relative to continuing employees was flat during this period.

We also calculate the average education level of unemployed workers relative to that of continuing employees by skill sector. In the presence of upskilling, based on selective employment we would expect the pool of unemployed workers in an occupation or skill sector to eventually fall behind continuing employees in terms of their education levels. Consistent with this prediction, Figure 8 shows that, among high-skill occupations where much of the permanent upskilling occurred, the education levels of unemployed individuals relative to continuing workers has been falling over time. In contrast, among middle-skill occupations where much of the temporary upskilling occurred, the relative education of unemployed workers relative to continuing workers fell and then increased, suggesting that these workers were shut out of jobs during the recession but then were able to gain employment when education requirements reverted back to their pre-recession levels. Finally, among low-skill occupations where most occupations experience no upskilling, the education level of the unemployed relative to continuing employees was flat during this period.

5. Labor Market Mismatch by Skill Sector

In the previous sections, we find that movements in the demand for and supply of skills vary across low-, middle-, and high-skill sectors. On the demand-side, education requirements increased across all occupations during the recession but stayed elevated only among high-skill occupations during the recovery, indicating a permanent increase over the course of the business
cycle. In contrast, middle-skill occupations exhibited more temporary upskilling versus low-skill occupations that experienced little to no upskilling during the recession and recovery periods. Comparing trends in requirements for actual categories of skill, the increase in baseline, specialized and software skills was larger for permanent versus temporary upskilling occupations, but significantly so only for software skills. High-skill occupations are more likely to require all three types of skills, but especially software skills that are specialized and employers are willing to pay workers with these skills a higher salary—even within six-digit occupations. On the supply-side, the education levels of new hires in the high-skill exceeded that of continuing employees while those of unemployed workers fell below, suggesting that job searchers may no longer be qualified for jobs within that sector. In contrast, this appeared to be a temporary phenomenon within the middle- and low-skill sectors. We hypothesize that structural increases in skill requirements that were concentrated within high-skill occupations may lead to greater persistence in labor market mismatch during the recovery compared to middle- and low-skill occupations. To test this, we follow Sahin et al. to construct separate mismatch indices for the low-, middle-, and high-skill sectors. Figure XX plots the basic index, $M_t$, and the productivity-adjusted index, $M_{xt}$, for 2007 through 2017 (the period for which the normalized BGT data is available).

As seen in Figure 9, several interesting patterns emerge. First, the degree of mismatch increases with the level of education required, suggesting that while more education makes workers more adaptive, it also makes them more specialized and the second effect seems to dominate. The mismatch index is highest for the high-skill occupations (0.2-0.25) and lowest for low-skill occupations (0.04 to 0.075). This confirms the notion that workers in occupations that require higher levels of education are likely to be more specialized and as a result, less substitutable
across occupational categories—even at the two-digit level. For example, an individual with a bachelor’s degree or higher working in the Architectural and Engineering occupational group is not likely to be able 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.

Moreover, changes over time in our measure of mismatch also vary by education group. Among high-skill occupations, mismatch has been increasing slightly since the recession, consistent with a permanent increase in skill requirements. In contrast, mismatch among middle-skill occupations fell (subject to slight fluctuations) during the recovery from 0.10 in 2010 to 0.07 in late 2017, as skill requirements receded during the recovery after the temporary run-up during the recession. Low-skill mismatch also decreased during the recovery, although to a lesser degree as fewer occupations in this skill sector had experienced any upskilling during the recession. These findings are consistent with other evidence from the literature. For example, job polarization has been characterized by the disappearance of routine and manual job opportunities in middle-skill occupations (Autor and Acemoglu 2011) particularly during recessions (Jaimovich and Siu 2014).

As a result, Figure 10 shows that the degree to which mismatch contributes to the decline in the unemployment rate also varies considerably by education group. Among high-skill occupations requiring a bachelor’s or greater, mismatch contributed upwards of 1.5 percentage points to the unemployment rate at the peak and remains elevated with only a partial decline since then. In contrast, mismatch among middle-skill occupations, contributed only 1 percentage point to the peak unemployment rate in 2010 and remains only slightly elevated from its pre-recession value as of 2017. Similarly, mismatch among low-skill occupations contributed more than 1.0 percentage point to unemployment in 2010 but as of 2017 contributes roughly the same amount (0.46 percentage point) as it did in 2007 (0.43 percentage point).
Table 4 summarizes the result of mismatch unemployment estimates across skill sectors. The contribution of declining occupational mismatch to the fall in unemployment between December 2007 and December 2017 is greater as we move from the lowest to the highest education category. In particular, among low-skill occupations, mismatch explains a little less than 0.46 percentage point (7 percent) of the 5.3 percentage point decline in the unemployment rate for that skill sector. In contrast, mismatch explains 0.52 (16 percent) out of the 3.3 percentage point decline in unemployment for middle-skill occupations and 0.66 (21 percent) out of a 3.16 percentage point fall in unemployment for the high-skill occupations. Also, between 2007 and 2015, the share of unemployment due to mismatch increased significantly from 31 percent to 47 percent among high—skill occupations while falling slightly for middle- and low-skill occupations.

How can these results be used to reconcile the puzzling observation in the literature that the Beveridge Curve appears to have shifted outwards but with little evidence of an increase in wages that would indicate a clear signal of mismatch in the labor market? Since the onset of the Great Recession, the Beveridge Curve has displayed a higher number of vacancies for a given level of unemployment, even as the labor market has recovered. Figure 11 draws the Beveridge Curve for each of our three skill sectors for the period 2007-2018 A notable distinction across skill sectors is that the slope of the Beveridge Curve is correlated with education such that high-skill occupations exhibit the steepest relationship and low-skill occupations have the flattest. In addition, most of the improvement in the Beveridge Curve has come from movements in the curve for low- and middle-skill occupations which 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, possibly
accounting for the persistent “wedge” that economists have observed in the aggregate Beveridge Curve during this period.

6. Conclusion and Policy Implications

In this paper, we test the hypothesis that changes in skill requirements within some occupations have reduced matching efficiency within some classes of jobs and contributed to the outward shift in the Beveridge Curve since 2007. We find evidence that supports this hypothesis from a few different sources. We describe movements in employer education requirements over time using a novel database of 87 million online job posting aggregated by Burning Glass Technologies. These data offer evidence of labor demand shifts among high-skilled occupation groups, which exhibited a permanent increase in the share of employers requiring a bachelor’s degree. Moreover, this increase in education requirements is correlated with concurrent increases in baseline, specialized, and software skills listed on job postings, suggesting a role for structural shifts associated with changes in technology or capital investment. For example, we observe large and continuous increases in the demand for software skills between 2007 and 2017 among the group of occupations that enacted permanent increases in education requirements since the Great Recession. Demand for software skills increased even among temporary upskillers, but in the latter case the increases occurred only between 2007 and 2010 rather than persisting over the entire period 2007 to 2017.

We also extend the framework developed by Sahin et al. (2014) to measure labor market mismatch, decomposing measures of mismatch by the education requirements of the underlying occupations. We find that between 2007 and 2017, high-skill occupations displayed higher mismatch rates than did either middle-skill or low-skill occupations, and high-skill occupations experienced a smaller decline in mismatch rates during the recent recovery. Adding to the story
we plot Beveridge curves separately by skill sector and we observe that high-skill occupations display reduced matching efficiency relative to 2007 whereas the Beveridge curves for both low-skill or middle-skill occupations are closer to their original (2007) positions as of 2018. Together the results suggest that the persistently lower matching efficiency among high-skill jobs—evidenced by the Beveridge curve for the high-skill sector—may be an artifact of increasing skill requirements within jobs, rather than the result of changes in the relative demand for occupations.

These findings have important implications for both the economics literature as well as labor market policy. Regarding the literature, our results demonstrate that equilibrium models where unemployed workers accumulate specific human capital and, in equilibrium, make explicit mobility decisions across distinct labor markets, can be chasing a moving target—at least among high-skilled occupations (Kambourov and Manovskii 2009; Alvarez and Shimer 2011; Carrillo-Tudela and Visscher 2013; and Wiczer 2013). Going forward, these frameworks can be modified to investigate the dynamics causing job seekers to search for work in the wrong sectors.

In terms of policy-making, the characteristics of occupations experiencing more permanent shifts in skill and education requirements can point to the potential structural forces underlying these observed trends. To that end, our findings can inform debates focused on workforce development strategies and related educational policies where decision making could benefit from the use of real-time labor market information on employer demands to provide guidance for both job placement as well as program development.

References


Mukoyama, Toshihiko, Christina Patterson, and Aysegul Sahin. 2014. "Job search behavior over the business cycle," Staff Reports 689, Federal Reserve Bank of New York.


Figure 1: Beveridge Curve (Revised as of July 2018, BLS)

Figure 2. Comparison of JOLTS and BGT Data Over Time

**Figure 3.** Trend in Share of Postings Requiring Various Levels of Education by Skill Group

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

Source: Authors’ calculations using data on online job vacancies from Burning Glass Technologies for 2007 and 2010-2017.

Note: Low-skill occupations are defined as those employing at least 40 percent of workers with a high school education or less according to the 2005-07 combined American Community Survey. High-skill occupations are defined as those employing at least 40 percent of workers with a Bachelor’s degree or greater according to the 2005-07 combines American Community Survey. Middle-skill occupations are all other occupations that have no clear plurality of low- or high-skill workers.
Figure 4. Trend in Share of Postings Requiring Various Types of Skills by Skill Group

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

Source: Authors’ calculations using data on online job vacancies from Burning Glass Technologies for 2007 and 2010-2017.

Note: Low-skill occupations are defined as those employing at least 40 percent of workers with a high school education or less according to the 2005-07 combined American Community Survey. High-skill occupations are defined as those employing at least 40 percent of workers with a Bachelor’s degree or greater according to the 2005-07 combines American Community Survey. Middle-skill occupations are all other occupations that have no clear plurality of low- or high-skill workers.
Figure 5. Initial versus Change in Share of Postings Requiring Baseline Skills, 2010-17

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

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

Note: See notes on Figure 4.
Figure 6. Initial versus Change in Share of Postings Requiring Specialized Skills, 2010-17

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

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

Note: See notes on Figure 4.
Figure 7. Initial versus Change in Share of Postings Requiring Software Skills, 2010-17

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

Source: Authors’ calculations using data on online job vacancies from Burning Glass Technologies for 2010 and 2017.
Note: See notes on Figure 4.
Figure 8. Relative Education Level of New Hires Versus Continuing Employees, 2005-17

A. Low-Skill Occupations

B. Middle-Skill Occupations

C. High-Skill Occupations

Source: Authors’ calculations using IPUMS-CPS data and IPUMS-ACS data.
Note: Low-skill occupations are defined as those employing at least 40 percent of workers with a high school education or less according to the 2005-07 combined American Community Survey. High-skill occupations are defined as those employing at least 40 percent of workers with a Bachelor’s degree or greater according to the 2005-07 combines American Community Survey. Middle-skill occupations are all other occupations that have no clear plurality of low- or high-skill workers.
Figure 9. Occupational Mismatch by Education Groups (2 Digit SOC Level)

A. High Skill Occupations

B. Middle Skill Occupations

C. Low Skill Occupations

Source: Authors’ calculations using occupation-level monthly job postings data from BGT and occupation-level monthly unemployment and labor force estimates from CPS covering January 2010 to December 2015.

Note: See the appendix for details.
Figure 10. Mismatch Unemployment by Education Groups (2 Digit SOC Level)

A. High Skill Occupations

B. Middle Skill Occupations

C. Low Skill Occupations

Source: Authors’ calculations using occupation-level monthly job postings data from BGT and occupation-level monthly unemployment and labor force estimates from CPS covering January 2010 to December 2015.

Note: See the appendix for details.
Figure 11. Beveridge Curve by Education Groups (2-Digit SOC Level)

A. High-Skill Occupations
B. Beveridge Curve, Middle-Skill Occupations
C. Beveridge Curve, Low-Skill Occupations

Source: Authors’ calculations using occupation-level monthly job postings data from BGT and occupation-level monthly unemployment and labor force estimates from CPS covering January 2010 to December 2015

Note: See the appendix for details.
Table 1. Share of 3-digit Occupations that were Permanent vs. Temporary Upskillers by Skill Group, 2007-17

<table>
<thead>
<tr>
<th></th>
<th>Percent of Occupations that Engaged in:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Permanent Upskilling</td>
</tr>
<tr>
<td>All Occupations</td>
<td>39.6%</td>
</tr>
<tr>
<td>Low-Skill Occupations</td>
<td>11.6%</td>
</tr>
<tr>
<td>Middle-Skill Occupations</td>
<td>40.0%</td>
</tr>
<tr>
<td>High-Skill Occupations</td>
<td>82.1%</td>
</tr>
</tbody>
</table>


Note: Permanent upskilling is defined as having a 5 percentage point or greater increase in share of postings requiring a BA or higher between 2007 and 2010 and having less than a 2.5 percentage point decrease in that share between 2010 and 2017. Temporary upskilling is defined as having a 5 percentage point or greater increase in share of employers requiring a BA or higher between 2007 and 2010 and having a 2.5 percentage point or greater decrease in that share between 2010 and 2017. No upskilling is defined as having less than a 5 percentage point increase in share of postings requiring a BA or higher. Shares are weighted by the occupation’s share of total employment as of 2006. Low-skill occupations are defined as those employing at least 40 percent of workers with a high school education or less according to the 2005-07 combined American Community Survey. High-skill occupations are defined as those employing at least 40 percent of workers with a Bachelor’s degree or greater according to the 2005-07 combines American Community Survey. Middle-skill occupations are all other occupations that have no clear plurality of low- or high-skill workers.
Table 2. Share of 3-digit Occupations by Skill Group that were Permanent vs. Temporary Upskillers, 2007-17

<table>
<thead>
<tr>
<th></th>
<th>Low-Skill</th>
<th>Mid-Skill</th>
<th>High-Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Occupations</td>
<td>47.3%</td>
<td>22.0%</td>
<td>30.8%</td>
</tr>
<tr>
<td>Permanent Upskillers</td>
<td>13.9%</td>
<td>22.2%</td>
<td>63.9%</td>
</tr>
<tr>
<td>Temporary Upskilling</td>
<td>47.1%</td>
<td>29.4%</td>
<td>23.5%</td>
</tr>
<tr>
<td>No Upskilling</td>
<td>79.0%</td>
<td>18.4%</td>
<td>2.6%</td>
</tr>
</tbody>
</table>


Note: Permanent upskilling is defined as having a 5 percentage point or greater increase in share of postings requiring a BA or higher between 2007 and 2010 and having less than a 2.5 percentage point decrease in that share between 2010 and 2017. Temporary upskilling is defined as having a 5 percentage point or greater increase in share of employers requiring a BA or higher between 2007 and 2010 and having a 2.5 percentage point or greater decrease in that share between 2010 and 2017. No upskilling is defined as having less than a 5 percentage point increase in share of postings requiring a BA or higher. Shares are weighted by the occupation's share of total employment as of 2006. Low-skill occupations are defined as those employing at least 40 percent of workers with a high school education or less according to the 2005-07 combined American Community Survey. High-skill occupations are defined as those employing at least 40 percent of workers with a Bachelor’s degree or greater according to the 2005-07 combines American Community Survey. Middle-skill occupations are all other occupations that have no clear plurality of low- or high-skill workers.
Table 3. Difference-in-Difference Analysis of Skill Requirements for Permanent versus Temporary Upskilling Occupations, 2007-2010-2017

A. Change in Skill Requirements during Recession: 2007 to 2010

<table>
<thead>
<tr>
<th></th>
<th>Share Requiring a BA</th>
<th></th>
<th></th>
<th>Diff</th>
<th>Diff-in-Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2007</td>
<td>2010</td>
<td>Diff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent Upskillers</td>
<td>19.49</td>
<td>37.29</td>
<td>17.80*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporary Upskillers</td>
<td>13.87</td>
<td>27.43</td>
<td>13.56*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>-5.61</td>
<td>-9.86</td>
<td></td>
<td></td>
<td>-4.24*</td>
</tr>
</tbody>
</table>

|                        | Share Requiring Baseline Skills |  |               |        |              |
|                        | 2007  | 2010  | Diff   |        |              |
| Permanent Upskillers   | 38.56 | 66.61 | 30.05*  |        |              |
| Temporary Upskillers   | 33.74 | 65.12 | 31.38*  |        |              |
| Diff                   | -4.82 | -3.49 |        |        | 1.33*        |

|                        | Share Requiring Specialized Skills |  |               |        |              |
|                        | 2007  | 2010  | Diff   |        |              |
| Permanent Upskillers   | 58.07 | 81.09 | 23.03*  |        |              |
| Temporary Upskillers   | 47.46 | 73.56 | 26.11*  |        |              |
| Diff                   | -10.61| -7.53 |        |        | 3.08         |

|                        | Share Requiring Software Skills |  |               |        |              |
|                        | 2007  | 2010  | Diff   |        |              |
| Permanent Upskillers   | 13.48 | 24.97 | 11.49*  |        |              |
| Temporary Upskillers   | 9.29  | 20.23 | 10.93*  |        |              |
| Diff                   | -4.19 | -4.74 |        |        | -0.56        |

|                        | Share Requiring Any Skills |  |               |        |              |
|                        | 2007  | 2010  | Diff   |        |              |
| Permanent Upskillers   | 65.62 | 88.11 | 22.49*  |        |              |
| Temporary Upskillers   | 51.92 | 80.83 | 28.92*  |        |              |
| Diff                   | -13.71| -7.28 |        |        | 6.43*        |

B. Change in Skill Requirements during Recovery: 2010 to 2017

<table>
<thead>
<tr>
<th></th>
<th>Share Requiring a BA</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2010</td>
<td>2017</td>
<td>Diff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent Upskillers</td>
<td>37.29</td>
<td>41.11</td>
<td>3.82*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporary Upskillers</td>
<td>27.43</td>
<td>14.88</td>
<td>-12.55*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>-9.86</td>
<td>-26.23</td>
<td></td>
<td></td>
<td>-16.37*</td>
</tr>
</tbody>
</table>

|                        | Share Requiring Baseline Skills |  |               |        |              |
|                        | 2010  | 2017  | Diff   |        |              |
| Permanent Upskillers   | 68.61 | 79.98 | 11.38* |        |              |
| Temporary Upskillers   | 65.12 | 76.03 | 10.91* |        |              |
| Diff                   | -3.49 | -3.96 |        |        | -0.47        |

|                        | Share Requiring Specialized Skills |  |               |        |              |
|                        | 2010  | 2017  | Diff   |        |              |
| Permanent Upskillers   | 81.09 | 93.39 | 12.30* |        |              |
| Temporary Upskillers   | 73.56 | 89.30 | 15.73* |        |              |
| Diff                   | -7.53 | -4.09 |        |        | -3.44        |

|                        | Share Requiring Software Skills |  |               |        |              |
|                        | 2010  | 2017  | Diff   |        |              |
| Permanent Upskillers   | 24.97 | 33.85 | 8.88*  |        |              |
| Temporary Upskillers   | 20.23 | 20.75 | 0.52   |        |              |
| Diff                   | -4.74 | -3.11 |        |        | -8.37*       |

|                        | Share Requiring Any Skills |  |               |        |              |
|                        | 2010  | 2017  | Diff   |        |              |
| Permanent Upskillers   | 88.11 | 100.00| 11.89* |        |              |
| Temporary Upskillers   | 80.83 | 100.00| 19.17* |        |              |
| Diff                   | -7.28 | 0.00  |        |        | 7.28*        |

Notes: See notes on Table 2.
Table 4. Change in Mismatch Unemployment by Skill Sector

<table>
<thead>
<tr>
<th>Skill Sector</th>
<th>$u_{10} - u_{10}^*$</th>
<th>$u_{10} - u_{10}^*$/$u_{10}$</th>
<th>$u_{15} - u_{15}^*$</th>
<th>$u_{15} - u_{15}^*$/$u_{15}$</th>
<th>$\Delta (u - u^*)$</th>
<th>$\Delta u$</th>
<th>$\Delta (u - u^*)/\Delta u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Skill Occupations</td>
<td>0.87</td>
<td>7</td>
<td>0.46</td>
<td>6</td>
<td>-.46 ppts</td>
<td>-5.38 ppts</td>
<td>7</td>
</tr>
<tr>
<td>Middle Skill Occupations</td>
<td>0.94</td>
<td>14</td>
<td>0.41</td>
<td>12</td>
<td>-.52 ppts</td>
<td>-3.33 ppts</td>
<td>16</td>
</tr>
<tr>
<td>High Skill Occupations</td>
<td>1.43</td>
<td>31</td>
<td>0.82</td>
<td>47</td>
<td>-.56 ppts</td>
<td>-3.16 ppts</td>
<td>21</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations using data from Burning Glass Technologies and the Current Population Survey. Note: $\Delta u$ varies by skill level. See the appendix for details.
A. Online Job Vacancy Data

1. Representativeness of Online Job Posting Data Compared to Surveys of Job Openings

Over the past two decades we have seen the development of surveys of job openings at both the federal and state levels to address the need for regular timely data on job openings, hires, and separations. At the federal level, the Job Openings and Labor Turnover Survey (JOLTS) is a monthly survey designed to serve as demand-side indicators of labor shortages at the national level. Prior to JOLTS, there was no economic indicator of the unmet demand for labor with which to assess the presence or extent of labor shortages in the United States. At the same time, about a dozen states also began collecting similar data on a local level, although at a lower frequency. Below we provide a description of both the JOLTS as well as the Minnesota Job Vacancy Survey and discuss their comparability to the newer sources of online job postings that have recently been made available to researchers.

1.1 Job Openings and Labor Turnover Survey (JOLTS)

Each month the JOLTS sample is comprised of approximately 16,000 businesses drawn from 8 million establishments represented in the Quarterly Census of Employment and Wages. The publicly available data provides a measure of labor demand across broad industry classifications at the national level or overall aggregate labor demand for the four Census divisions. JOLTS does not provide measures by occupation nor at more refined geographies.

\[\text{\footnotesize{14 See https://www.bls.gov/jlt/jltopen.htm for more details.}}\]
JOLTS collects data on total employment, job openings, hires, quits, layoffs and discharges, and other separations. JOLTS does not collect information on the skill requirements (e.g. education and experience). Job openings are defined as all positions that are open (not filled) on the last business day of the month. A job is "open" only if it meets all three of the following conditions:

- A specific position exists and there is work available for that position. The position can be full-time or part-time, and it can be permanent, short-term, or seasonal, and
- The job could start within 30 days, whether or not the establishment finds a suitable candidate during that time, and
- There is active recruiting for workers from outside the establishment location that has the opening, meaning that the establishment is taking steps to fill a position. It may include advertising in newspapers, on television, or on radio; posting Internet notices; posting "help wanted" signs; networking with colleagues or making "word of mouth" announcements; accepting applications; interviewing candidates; contacting employment agencies; or soliciting employees at job fairs, state or local employment offices, or similar sources.

The JOLTS measure of job openings does not include

- Positions open only to internal transfers, promotions or demotions, or recall from layoffs
- Openings for positions with start dates more than 30 days in the future
- Positions for which employees have been hired, but the employees have not yet reported for work
- Positions to be filled by employees of temporary help agencies, employee leasing companies, outside contractors, or consultants.
1.2 Minnesota Job Vacancy Survey

The Minnesota Job Vacancy Survey is one of twelve state job vacancy surveys conducted in the United States, collected biannually since 2001.Using the QCEW sampling frame, it is designed to estimate hiring demand and job vacancy characteristics by industry and occupation. Information is gathered through the survey of a stratified sample of about 10,000 firms in 13 regions of Minnesota. Firms excluded from the sampling process include private households, personnel service industry establishments and businesses with no employees. Job vacancies reserved for contract consultants, employees of contractors and others not considered employees of surveyed firms are excluded.

Job vacancies are defined as positions that are currently open-for-hire at the time of the survey based on a set of questions similar to those employed by the JOLTS. The survey also asks employers about the minimum requirements for education and experience for each vacancy. The survey only reports three distinct categories for experience: no work experience, some work experience, and related work experience. The constructed dependent variable for experience identifies whether the job requires related experience only.

1.3 Comparison to Online Job Posting Data

Over the past two decades, online vacancy data have been used by a number of researchers to study labor market dynamics (e.g., Sahin et al. 2014, Marinescu and Wolthoff 2013, Lazear and Spletzer 2012, Faberman and Mazumder 2012, Rothwell 2012, Bagues and Labini 2009, Kuhn and Skuterud 2004, Gautier, van der Berg, van Ours, and Ridder 2002). The advantage of using online vacancy data is that they allow analysis at a greater frequency and at

---

more refined geographies than traditional employer surveys, such as the Job Opening and Labor Turnover Survey (JOLTS). This is because the data are constructed from measures that are collected by software that parses the text contained in millions of job ads posted online on a daily basis. One potential drawback is that online vacancy data only capture vacancies posted on the Internet, and may not be representative of the universe of job postings if vacancies from certain industries and occupations are less likely to be posted electronically. However, a recent report from Georgetown University estimates that between 60 and 70 percent of job postings are now posted online (Carnevale, Jayasundera, and Repnikov 2014). Other researchers have shown that online job ads exhibit similar trends and are closely correlated with employer surveys over time as well as across industries and occupations (Templin and Hirsch 2013, Ganong 2014).

There is also a concern that the use of online job postings has changed over the sample period, although again there is reason to believe that this is less of an issue during the time period of our analysis compared to earlier years. The first online job listings were posted on Usenet, CareerMosaic, and Monster during 1990-1994. Between 1995 and 1997, additional job boards were launched (e.g. Craigslist) and newspaper job listings began to appear online. Major changes took place in the years 1998 and 1999 as the job boards industry consolidated and a few key players emerged (e.g. Monster, Career Builder, and Jobsonline). After the dot-com bust, niche job boards proliferated between 2000 and 2002 for marketing, medical, sales, and accounting jobs. Between 2003-2007, the industry matured and experienced significant growth with the launch of LinkedIn.¹⁶ All to say that by 2007, the starting point of our analysis, online job posting was quite prevalent among employers.

Moreover, aggregating job posting data is not a new endeavor either. Two companies, Jungle and Careercast, began scraping and aggregating online job postings in the late 1990s. By the early 2000s, aggregators such as Top USA Jobs, Indeed, and Simply Hired emerged. By 2008, The Conference Board discontinued its-long-running, print-based Help-Wanted Advertising Index, after having begun a Help-Wanted Online Index (HWOL) in 2005 based on web scraping process similar to that of Burning Glass Technologies.\(^17\)

2. **Representativeness of Burning Glass Technologies Data**

The main source of data used in this paper is collected by Burning Glass Technologies (BGT), one of the leading vendors of online job posting data.\(^18\) BGT collects detailed information on the more than seven million current online job openings daily from over 40,000 sources including job boards, newspapers, government agencies, and employer sites. The collection process employed by BGT provides a robust representation of online hiring, including job activity 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 text parse each job ad into usable data elements. The process follows a fixed schedule, “spidering” a pre-determined basket of websites that is carefully monitored and updated to include the most current and complete set of online postings. BGT has developed algorithms to eliminate duplicate ads for the same job posted on both an employer website as well as a large job board by identifying a series of identically parsed variables across job ads such as location, employer, and job-title. In addition, to avoid large fluctuations over time, BGT places more weight on large job boards than

\(^{17}\) See [https://www.conference-board.org/data/helpwantedonline.cfm](https://www.conference-board.org/data/helpwantedonline.cfm).

individual employer sites which are updated less frequently. In addition, their Labor/Insight analytical tool 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.

BGT mines over seventy job characteristics from free-text job postings including employer name, location, job-title, occupation, years of experience requested and level of education required or preferred by the employer. 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 filled.

BGT provides snapshots of the data in which vacancies are reported on a monthly basis and are pooled over the year without duplication. As such, this data is unique in allowing geographical analysis of occupation-level labor demand by education level and experience level. The data are available for detailed occupation by Standard Occupation Code (SOC) down to the six-digit level. Over the three years that we examine (2007, 2010, and 2012), the BGT data represent roughly 66.8 million vacancies during this period.

In addition to the concerns regarding the representativeness of online job vacancy data in general, it is also possible that the BGT data is not representative of the universe of online job postings. Coverage depends on the number of sites BGT scrapes as well as proprietary algorithms to recognize and eliminate duplicates as well as to accurately code information in the postings. However, since we originally began using the BGT data in 2014, a number of other authors have also conducted research using these data, validating its robustness in the process. For example, Carnevale, Jayasundera, and Repnikov (2014) show that the occupation-industry composition of the BGT data are similar to that of the Conference Board's HWOL. Moreover, the authors audited a sample of job postings in the BGT database and compared them to the
actual text of the postings. They concluded that the BGT coding for occupation, education, experience was accurate at least 80 percent of the time. Given that BGT has made improvements to their coding algorithms since the version of the data studied by Carnevale, Jayasundera, and Repnikov (2014), it’s likely that the accuracy of our sample is even higher.

Finally, we note that the chief concern regarding the BGT data is about the range of external validity for our results, as opposed to the introduction of omitted variable bias. While we use the BGT data to construct the dependent variables in our analysis, we are careful to include controls for changes over time and composition in our specifications while also using an exogenous shock to supply with returning veterans. Nonetheless, below we discuss the representativeness of online job vacancy data in general, and the BGT data specifically, and investigate whether the representativeness has changed during the time period of our analysis. This includes a comprehensive comparison to other sources of job vacancy data (e.g., JOLTS and the Minnesota Job Vacancy Survey), and an overview of other research validating this new data source.

2.1 Representativeness of BGT Data Over Time

Previous studies have not used the BGT data to analyze changes over time in skill requirements and as such we discuss the consistency and representativeness of the data during the time period in our study in this section. In terms of consistency, while we do raise the concern that the nature of the BGT data may have changed over time, this is less of an issue with the version of the data that BGT has provided to us. Although BGT regularly improves the algorithms that are used to clean job-titles and employer names, categorize job postings by occupation and industry and extract additional information from text such as location and other skills requested, they apply these algorithms retroactively to the complete historical database of
postings. Thus we are not concerned about changes over time in how the data are constructed. In addition, we provide specifications that draw on a panel of employers and job-titles which eliminates any changes in the composition of the sample that we study.

With regards to representativeness, we provide several basic robustness checks below during the time period in our study. First, we compare the overall trend in the number of job postings during the Great Recession for BGT versus JOLTS and the Minnesota Job Vacancy Survey. In Table A.1 below, the percent change in the number of postings over time are very similar showing a roughly 30 percent decline between 2007 and 2010 and a similar increase between 2010 and 2012. Note that the number of postings in the BGT data is much lower than that of the JOLTS, similar to other online job posting data such as HWOL. In part, this is because JOLTS asks employers how many job openings are available whereas the online vacancy data only capture one posting which may represent multiple openings. In addition, some occupations and industries are not fully captured by the BGT such as those that are likely to employ spot labor (e.g. construction) or those that source candidates through channels other than online postings (e.g. public service jobs).

In particular, one might be concerned that the distribution of postings by industry and occupation in the BGT data might not be representative or might have changed over time. Figure A1 plots the distribution of BGT postings across major industry groups for 2007, sorted from largest to smallest, as well as the distribution of job vacancies in JOLTS. Despite the differences in the sampling of the two sources, the industry distributions are quite similar—the deviation of BGT from JOLTS is less than 10 percentage points across all major categories. The largest deviations show that the BGT data tend to over-represent industries such as finance and

---

19 The Minnesota Job Vacancy Survey is one of twelve state job vacancy surveys conducted in the United States, collected biannually since 2001. See the Data Appendix for further details.
insurance, health care and social assistance, and educational services and under-represent food
services, public administration/government, and construction. Again, most 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
JOLTS 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 JOLTS job vacancies in the same year. For example, finance
and insurance is shown on the far right, at roughly 7 percentage points overrepresentation in
BGT compared to JOLTS in 2007 and accommodation and food services is on the far left, at
roughly 7 percentage points underrepresented in BGT versus JOLTS for that same year. On the
y-axis we plot the same deviation for 2012. We also plot the 45-degree line as a benchmark: if
representativeness of the BGT data, relative to the JOLTS, remained constant over time, all
markers should line up on the 45-degree line.

The figure shows that changes in representativeness over this time period are very small
with most of the markers lying close to the 45 degree line. To the extent that changes did occur,
there is a tendency for them to have been in the direction of closer representativeness to the
JOLTS. For example, finance and insurance was only over-represented by 4 percentage points in
2012 and accommodation and food services was only under-represented by 4.5 percentage
points. Yet for all of these industry groupings, the change in representation was over time was
quite small.

The primary advantage of the BGT data over the JOLTS is that they allow us to
categorize jobs by occupation at a detailed level. We thus also compare the occupational
distribution of BGT job postings to the Minnesota Job Vacancy Survey. Figure A3 plots the
distribution of postings across broad occupation groups, sorted from largest to smallest, for the
distribution of BGT online postings and MN JVS job vacancies. As with the JOLTS
comparisons, the occupational distributions are quite similar despite the differences in sampling.
The deviation of BGT from the MN JVS is again less than 10 percentage points across all major
categories. The largest deviations show that the BGT data tend to over-represent higher-skilled
occupations such as computer and mathematical (+9 percentage points), management (+7
percentage points), and business and financial (+6 percentage points) while under-representing
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.\textsuperscript{20}

Again, we are concerned with whether the representativeness 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 as we did earlier with the JOLTS. As with the industry
comparisons, changes in representativeness of occupations over this time period are very small
with most of the markers lying close to the 45 degree line. To the extent that changes did occur,
some of the differences narrowed over time. For example, business and financial occupations
were over-represented by +6 percentage point in 2007 but by only +5 percentage points in 2012.
Similarly, healthcare support occupations were under-represented by -5 percentage point in 2007
but by only -3 percentage points in 2012. Yet other occupations such as sales and healthcare

\textsuperscript{20} Rothwell (2014) also compares the occupational distributions from an extract of BGT to those from state vacancy
surveys for select metropolitan areas for which data are available. He finds that computer, management, and
business occupations are overrepresented relative to the state vacancy surveys, while health care support,
transportation, maintenance, sales, and food service workers are underrepresented.
practitioners and technical positions became less well-represented although again the changes were quite small (a two percentage point difference or less).

3. Remaining Questions

Finally, we acknowledge that there are still some remaining questions about how to think about online postings that are difficult to answer. For example, it’s true that we still don’t know that much about how many jobs are actually filled based on recruiting via the online postings. However, for our purposes, we are agnostic about whether the job ultimately is filled by a candidate that applies online versus some other channel (e.g. internal or networking). Instead, we take the job posting as an indication that a given employer is looking to fill at least one position and what the preferred qualifications are for hiring a job candidate from any source.

That said, data collected by SilkRoad technology, a global provider of social talent management solutions, indicates that online advertising has become the primary source of hiring for firms. SilkRoad teamed up with more than 700 of its customers to uncover which recruitment methods yield the largest number of interviews and hires. Rather than surveying its clients, SilkRoad collected primary data about them from OpenHire, its applicant tracking system, analyzing 222,308 job postings, 9.3 million applications, 147,440 interviews and 94,155 hires. Of the external recruitment marketing sources—which led to half of all interviews and about a third (37 percent) of all hires—online sources were by far the most effective. Search engines and job boards produced 94 percent of the interviews and 86 percent of the hires among those who utilized external sources. However, within the online category there has been a shift over the past several years. When SilkRoad first published their Top Sources of Hire report in 2012, job search engines accounted for only 40 percent of interviews and hires, with job boards

accounting for the other 60 percent. As of 2016, those shares were reversed with job search engines now accounting for 62 percent of hires.

In addition, a 2015 survey by the Pew Research Center found that online employment resources now rival personal and professional networks as a top source of job information for Americans who are looking for work. Using a nationally representative sample of 2,001 U.S. adults ages 18 and older, roughly one-third of those surveyed had looked for a new job in the last two years, and 79% of these job seekers utilized online resources in their most recent search for employment. That is higher than the proportion who made use of close personal connections (66%) or professional contacts (63%) and more than twice the proportion who utilized employment agencies, print advertisements, or jobs fairs and other events.22

Finally, there is at least some suggestive evidence that the BGT occupational distributions reflect that of new hires, implying that there is a reasonable connection between online job postings and employment. In a forthcoming study in the American Economic Review, Hershbein and Kahn (2017) validate this aspect of the BGT data by comparing the occupational distribution of BGT versus new jobs 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 health care 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.”

Figure A1. Industry Distribution of Job Postings: BGT versus JOLTS, 2007.

Figure A2. Industry Distribution of Job Postings: BGT versus JOLTS, 2007 and 2012.


Note: The x-axis is the BG 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.

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


Note: The x-axis is the BG 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>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>Percent Change</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2007</td>
<td>2010</td>
<td>2012</td>
<td>2007-10</td>
</tr>
<tr>
<td><strong>Total U.S.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOLTS</td>
<td>4,477.5</td>
<td>2,852.6</td>
<td>3,671.9</td>
<td>-36.3</td>
</tr>
<tr>
<td>BGT</td>
<td>1,060.4</td>
<td>786.7</td>
<td>1,057.6</td>
<td>-25.8</td>
</tr>
<tr>
<td><strong>Minnesota</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MN JVS</td>
<td>56.6</td>
<td>37.6</td>
<td>60.9</td>
<td>-33.5</td>
</tr>
<tr>
<td>BGT</td>
<td>56.0</td>
<td>36.5</td>
<td>58.2</td>
<td>-34.9</td>
</tr>
</tbody>
</table>

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 JOLTS are for average monthly job postings over the year.
B. Estimation of Mismatch Indices

In order to compute the mismatch indices, we estimate the parameters of the model described in the main text (equation 1) for the vacancy elasticity, the market-specific matching efficiencies $\varphi_{it}$ and the aggregate level matching efficiency $\Phi_t$. To do so we follow the methodology described in the online appendix to Sahin et al. (2014). In particular, using the notation defined in the main text, and letting $t$ denote year, to estimate aggregate level market efficiencies we run the following OLS regression:

$$\log \left( \frac{h_t}{u_t} \right) = \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \delta \log \left( \frac{v_t}{u_t} \right) + \epsilon_t$$

(4)

In the above, the coefficient on the ratio of vacancies to unemployed, $\delta$, captures aggregate matching efficiency. To estimate occupation-specific matching efficiencies we run analogous regressions using occupation-specific hires, vacancies, and unemployment numbers.

1.2. Unemployment Mismatch

To calculate the counterfactual unemployment level that is the level of unemployment in the absence of the sectoral/occupational labor mismatch, we have used the Shimer (2005) methodology. Let’s assume that all unemployed workers find a job with a probability of $F_t \in [0,1]$ and also employed individuals lose their jobs with probability of $S_t \in [0,1]$ in each period. Losing the job in each period happens according to a Poisson process with the arrival rate $s_t \equiv -\log(1 - S_t)$. Similarly, finding a job happens according to a Poisson process with the arrival rate of $f_t \equiv -\log(1 - F_t)$. 
The number of unemployed at date $t + 1$ can be solved as:

$$u_{t+1} = (1 - F_t)u_t + u^s_{t+1}$$  \hspace{1cm} (5)$$

Where $u^s_{t+1}$ denotes the short-term unemployed workers. Inverting the above formula, we can retrieve probability of job finding as a function of unemployment and short-term unemployment:

$$F_t = 1 - \frac{u_{t+1} - u^s_{t+1}}{u_t}$$  \hspace{1cm} (6)$$

Shimer (2005) solves the unemployment and short-term unemployment differential equations to obtain an implicit expression for the separation probability:

$$u_{t+1} = (1 - e^{-f_t - \sigma_t})s_t \frac{l_t + e^{-f_t - s_t}u_t}{f_t + s_t}$$  \hspace{1cm} (7)$$

Where $l_t \equiv u_t + e_t$ is the size of the labor force during period $t$.

Job finding rate in the absence of mismatch in each period is calculated using the following formula:

$$f^{ef}_t = \frac{f_t}{1 - M_t \left( \frac{u_t}{u^{ef}_t} \right)^a}$$  \hspace{1cm} (8)$$

Where $M_t$ is the estimated mismatch measure in each period. Finally, the level of unemployment in the absence of mismatch is$^{23}$.

$^{23}$ It is worth noting that the steady state of unemployment rate during each period could be derived as a function of job finding rate and job separation rate. In the steady state situation we have:

$$s_t, e_t = u_t f_t$$

Therefore, the steady state rate of unemployment is:
2. Data

We measure mismatch at both the industry and occupation levels. To construct the baseline mismatch index $M_t$ requires sectoral data on vacancies, unemployment, and the vacancy elasticity of the matching function. Data on vacancies and vacancy shares come from both the JOLTS (industry) and BGT (occupation). Data on unemployment comes from the CPS.

To construct the other mismatch indices requires additional data on market-specific matching efficiency parameters for the $M_{pt}$ index, and information on productive efficiency (productivity and separation rates) by sector for the $M_{xt}$ index and its corresponding counterfactual. To derive market-specific matching efficiencies and vacancy elasticities we will also require data on hires to estimate the appropriate matching functions.

\[
 u_{t+1}^{cf} = \left(1 - e^{-f_{t}^{cf} - s_{t}}\right)s_{t} + e^{-f_{t}^{cf} - s_{t}}u_{t}^{cf}
\]

The initial values of job finding rate and unemployment in the absence of mismatch are calculated as follows:

\[
 \frac{u_{t}}{l_{t}} = \frac{s_{t}}{s_{t} + f_{t}}
\]

and

\[
 f_{1}^{cf} = \frac{f_{1}}{1 - M_{t}}
\]

\[
 u_{1}^{cf} = l_{1} \cdot \frac{s_{1}}{s_{1} + f_{1}^{cf}}
\]
To compute the measures of labor mismatch based on the method that we explained in section two, we need detailed information on unemployment, vacancies, flows in and out of labor force, productivity rates, destruction rates, etc. We gather this information at the industry and occupational level.24

2.1. Vacancy Measures

2.1.1 Using the JOLTS data to Measure Baseline Industry Trends

The Bureau of Labor Statistics (BLS) collects and compiles Job Opening and Labor Turnover Survey (JOLTS) data monthly from a sample of roughly 16000 nonfarm establishments since December 2000. Data are collected for total employment, job openings, hires, quits, layoffs and discharges, other separations, and total separations. Detailed information on the number of vacant jobs25 and hires26 are available by the industry classification. JOLTS data provides 17 industrial classification which is comparable with two digits North American Industry Classification System (NAICS).27

24 "Industry" refers to the work setting and economic sector, while "occupation" relates to the worker's specific technical function.

25 According to JOLTS documentation vacancy definition is as follows: “Job openings information is collected for the last business day of the reference month. A job opening requires that: 1) a specific position exists and there is work available for that position, 2) work could start within 30 days whether or not the employer found a suitable candidate, and 3) the employer is actively recruiting from outside the establishment to fill the position.”

26 According to JOLTS documentation vacancy definition is as follows: “The hires level is the total number of additions to the payroll occurring at any time during the reference month, including both new and rehired employees, full-time and part-time, permanent, short-term and seasonal employees, employees recalled to the location after a layoff lasting more than 7 days, on-call or intermittent employees who returned to work after having been formally separated, and transfers from other locations. The hires count does not include transfers or promotions within the reporting site, employees returning from strike, employees of temporary help agencies or employee leasing companies, outside contractors, or consultants. The hires rate is computed by dividing the number of hires by employment and multiplying that quotient by 100.”

27 A more detailed discussion of JOLTS concepts and methodology is available online at www.bls.gov/opub/hom/pdf/homch18.pdf. Data files are downloaded from http://download.bls.gov/pub/time_series/jt/. Also, for detailed information about industry classification in JOLTS, see http://www.bls.gov/jlt/jlnaics.htm
Even though JOLTS estimates are survey based and therefore are subject to both sampling and non-sampling errors meaning that there is a chance that sample estimates may differ from the “true” population values that they present, they are still well known and widely used and more importantly designed to be comparable the supply side data that we use for our analysis that is unemployment estimates from Current Population Survey. However, we are not able to study structural imbalance at the occupational level and further analysis of labor mismatch by skill level using JOLTS. So instead we make use of the BGT vacancy measures for the occupation level analysis.

2.1.2 Using BGT Data to Measure Occupation Trends by Educational Groups

The other source of vacancy data used in this paper is collected by Burning Glass Technologies (BGT), one of the leading vendors of online job posting data. BGT collects detailed information on the more than seven million current online job openings daily from over 40,000 sources including job boards, newspapers, government agencies, and employer sites. 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 text parse each job ad into usable data elements. BGT mines over seventy job characteristics from free-text job postings including employer name, location, job title, occupation, years of experience requested and level of education required or preferred by the employer, as well as other dimensions of skill.

The collection process employed by BGT provides a robust representation of hiring, including job activity posted by small employers. The process follows a fixed schedule,

---

30 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 filled.
“spidering” a pre-determined basket of websites that is carefully monitored and updated to include the most current and complete set of online postings. BGT has developed algorithms to eliminate duplicate ads for the same job posted on both an employer website as well as a large job board by identifying a series of identically parsed variables across job ads such as location, employer, and job title. In addition, to avoid large fluctuations over time, BGT places more weight on large job boards than individual employer sites which are updated less frequently.\textsuperscript{31}

In the database provided by BGT, a snapshot of vacancies is reported on a monthly basis and are pooled over the year without duplication. This data is unique in allowing geographical analysis of occupation-level labor demand for a variety of skills including education and experience over time. Using the entire universe of job vacancies collected by BGT, allows us to expand the labor mismatch analysis to occupational level. The data are available for detailed occupation by Standard Occupation Code (SOC) down to the six-digit level for 2010 through end of 2015.\textsuperscript{32}

It should be noted that although Burning Glass Technologies consistently applies the same filtering and de-duplication algorithm across years, even retroactively as improvement are made, the number of sources scraped may have evolved over time. Figure 2 plots JOLTS vacancies and BGT ads at the national level. The total count of active vacancies in BGT is below that in JOLTS although the correlation between the two series is quite strong at about 0.85. To the extent that the trend in online vacancies is similar to that of JOLTS across sectors, our calculations should not be affected. Yet, there are also specific occupations which are

\textsuperscript{31} 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.

\textsuperscript{32} We have aggregated the data to the two-digit occupational categories using the appropriate mappings from the 2010 SOC codes.
underrepresented in all on-line job posting data. For example, construction jobs are not typically posted online. We utilize a re-weighting scheme introduced by BGT that adjust the total number of postings at the industry level by the number of monthly vacancies from JOLTS. It also utilizes other reliable sources of labor data such as quarterly workforce indicator (QWI), occupational employment statistics (OES), etc. to adjust the number of postings in cases that the number of aggregated BGT postings substantially differs from national trend. The reweighting process produces a BGT vacancy data series that is consistent and comparable with the JOLTS and other vacancy data series that use a similar weighting methodology such as HWOL.

Using the weighted BGT series as our starting point, we then make use of the detailed occupation and skill measures contained in the BGT data to construct mismatch indices for two-digit occupations by education level. Specifically, we use the education requirements contained in the BGT data to calculate the education distribution each year by six-digit Standard Occupational Classification (SOC) categories and classify six digit occupations as high-, middle-, and low-skill. We then aggregate the number of vacancies in each skill category up to the two digit levels to match up with the supply side measures in our mismatch analysis. This method yields separate vacancy measures by skill level at the two digit occupation classifications that consist only of those detailed occupations within that occupation grouping that meet those specific education criteria, rather than labeling the whole occupation grouping into one of those

33 High-skill occupations are those where at least 40 percent of the vacancies require a Bachelor’s degree or higher, low-skill occupations are those where at least 40 percent of the vacancies require a High School degree or less, and middle-skill occupations are those where the education requirements lie somewhere in between or require an Associate’s degree or some college.
buckets.\textsuperscript{34} We follow the same procedure to construct our supply-side measures using the Current Population Survey in the next section.

2.2 Constructing Other Variables: Unemployment, Hiring and Job Destruction Rates, and Productivity Measures

We use Current Population Survey (CPS) basic monthly data and applied person-level weights to estimate the monthly aggregate unemployment rates from January 2001 to December 2015.\textsuperscript{35} We construct estimates of unemployment and labor force counts by skill level for both industries as well as two-digit occupations. We also calculate aggregate unemployment rates for the same 17 industry category that we have the vacancy information from JOLTS to be able to replicate the results in Sahin et al. 2014 for comparison purposes.\textsuperscript{36}

We estimate the job finding rate for individuals who are surveyed in adjacent months of the CPS and use this proxy to calculate estimates of hires at the two-digit level for the occupation analysis. We use a similar strategy to calculate job destruction rate proxies at the occupational level. We use direct measure of hires and separations from the JOLTS data to calculate hiring and job destruction rates at the industry level.

Productivity measures at the industry level are calculated based on value added by employment level. Information on value added is gathered from the Business Employment

\textsuperscript{34} For example, vacancies within the two-digit Management occupation grouping (SOC 2010 = 11) would be allocated across high-, middle-, and low-skill levels according to the educational distribution for of the six digit occupations within it.

\textsuperscript{35} We utilize IPUMS CPS datasets for the supply side of our analysis. Please see https://cps.ipums.org/cps/ for detailed documentation on variables.

\textsuperscript{36} We have created crosswalk to map 17 JOLTS industries to NAICS categories and Census industry categories which is used in the CPS data. Appendix 1 covers some of the mappings used in the analysis. In the CPS for persons who were employed at the time of the survey, IND relates to the industrial sector in which the respondent worked during the preceding week. For unemployed persons and those not currently in the labor force, IND characterizes the industrial sector of the respondent’s most recent job. The CPS interviewer collected information by asking what kind of work the person was doing, and Census Bureau staff coded the information into the CPS or census industrial classification.
Dynamics (BED) database.\textsuperscript{37} Data on employment levels is constructed from the Establishment Survey for the 17 industry categories in JOLTS. We use median wage data as the proxy for occupational productivity using Occupational Employment Statistics (OES) Survey.

\textsuperscript{37} https://www.bls.gov/bdm/
## Table B.1 Industry Crosswalk

<table>
<thead>
<tr>
<th>Industry Categories</th>
<th>JOLTS</th>
<th>NAICS</th>
<th>Census 2003 to 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining and Logging</td>
<td>NAICS 1133—Logging, Sector 21—Mining</td>
<td>370-490,270</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>Sector 23—Construction</td>
<td>770</td>
<td></td>
</tr>
<tr>
<td>Durable Goods Manufacturing</td>
<td>NAICS 321, 327, Sector 33</td>
<td>2470–3990</td>
<td></td>
</tr>
<tr>
<td>Nondurable Goods Manufacturing</td>
<td>Sector 31, NAICS 322, 323, 324, 325, 326</td>
<td>1070–2390</td>
<td></td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>Sector 42—Wholesale Trade</td>
<td>4070–4590</td>
<td></td>
</tr>
<tr>
<td>Retail Trade</td>
<td>Sectors 44 and 45—Retail Trade</td>
<td>4670–5790</td>
<td></td>
</tr>
<tr>
<td>Transportation, Warehousing, and Utilities</td>
<td>Sectors 48 and 49—Transportation and Warehousing, Sector 22-</td>
<td>6070-6390, 570–</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>Sector 51—Information</td>
<td>6470–6780</td>
<td></td>
</tr>
<tr>
<td>Finance and Insurance</td>
<td>Sector 52—Finance and Insurance</td>
<td>6870–6990</td>
<td></td>
</tr>
<tr>
<td>Real Estate and Rental and Leasing</td>
<td>Sector 53—Real Estate and Rental and Leasing</td>
<td>7070-7190</td>
<td></td>
</tr>
<tr>
<td>Professional and Business Services</td>
<td>Sector 54, 55,56—Professional, Scientific, and Technical Services</td>
<td>7270–7790</td>
<td></td>
</tr>
<tr>
<td>Educational Services</td>
<td>Sector 61—Educational Services</td>
<td>7860–7890</td>
<td></td>
</tr>
<tr>
<td>Health Care and Social Assistance</td>
<td>Sector 62—Health Care and Social Assistance</td>
<td>7970–8470</td>
<td></td>
</tr>
<tr>
<td>Arts, Entertainment, and Recreation</td>
<td>Sector 71—Arts, Entertainment, and Recreation</td>
<td>8560–8590</td>
<td></td>
</tr>
<tr>
<td>Accommodation and Food Services</td>
<td>Sector 72—Accommodation and Food Services</td>
<td>8660–8690</td>
<td></td>
</tr>
<tr>
<td>Other Services</td>
<td>Sector 81—Other Services, except Public Administration</td>
<td>8770–9290</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>Sector 61 (public ownership)—State and Local Government Education</td>
<td>9370–9590</td>
<td></td>
</tr>
</tbody>
</table>
Figure B.1 Industry Level Correlation

Source: Authors’ calculations using disaggregated industry-level vacancy data extracted from JOLTS and industry-level monthly unemployment and labor force estimates from CPS covering January 2001 to December 2015.
Figure B.2 Industry Level Mismatch Indices and Unemployment Rates

A. Mismatch Indices

B. Unemployment Rates

Source: Authors’ calculations using disaggregated industry-level vacancy data extracted from JOLTS and industry-level monthly unemployment and labor force estimates from CPS covering January 2001 to December 2015
Figure B.3 Occupational Level Correlation

Source: Authors' calculations using occupation-level monthly job postings data from BGT and occupation-level monthly unemployment and labor force estimates from CPS covering January 2010 to December 2015.
Figure B.4 Two Digit Occupation-Level Mismatch Indices and Unemployment Rates

A. Mismatch Indices

B. Unemployment Rates

Source: Authors’ calculations using occupation-level monthly job postings data from BGT and occupation-level monthly unemployment and labor force estimates from CPS covering January 2010 to December 2015